VOLUME III FEBRUARY, 1912 NUMBER 2

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

A TENTATIVE REVISION AND EXTENSION OF THE BINET-SIMON MEASURING SCALE OF INTELLIGENCE.

LEWIS M. TERMAN,
Assistant Professor of Education, Stanford University,

AND

MR. H. G. CHILDS,
Assisted in the Tests by Katherin Romer Kip and Edith M. Bushnell.

PART I.[1]

INTRODUCTION.

In the light of the numerous recent studies of the progress of children through the grades it would appear that the time had arrived to determine more definitely than has yet been attempted the qualitative nature of the phenomena of retardation and acceleration. This determination should proceed along two lines. It should ascertain as far as possible the native intellectual endowment of each pupil, and it should undertake to establish with reference to acceptable standards his exact pedagogical status.

Though the latter task is by far the easier—is, in fact, quite feasible for most kinds of school work—we have barely set about its accomplishment. Except in handwriting and arithmetic, there are no generally acceptable standards of school

[1]This is the first of a series of articles. Part II will set forth the results of additional tests and Part III will give the author's revision and extension of the scale.

performance which enable us to make a reliable comparison of different pupils and schools. The average superintendent of schools accepts on faith the accomplishments of his own system and looks with suspicion on those of other schools, as is proved by the retardation which pupils usually suffer when they change from one school system to another. Such chaos cannot continue indefinitely now that the standarization has begun, and it does not seem unreasonable to expect that within a few years we shall have measures which will permit the expert determination of the efficiency of a school's efforts in most branches of the curriculum.

Measurements of intelligence are far more difficult to make, but even more important than those of pedagogical status. The degree of intelligence is a factor in every performance. Unfortunately, however, its influence cannot be isolated from that of training. Intelligence cannot work without ideas, and the quality and number of the latter are largely dependent upon accidents of environment. In our efforts to measure intelligence, therefore, we always measure intelligence plus a certain kind and amount of training. The only way to make sure that our measurement will arrive at any approximation to the former is to choose for our tests types of performance which will not be too greatly influenced by such differences in training and experience as ordinarily exist among the children of civilized people living under fairly uniform conditions of home life and educational advantages. While the social inheritance cannot, of course, be wholly identical for any two individuals, however similar their environment, there is nevertheless a limited social inheritance common to practically all who are possessed of a mentality sufficiently normal to make it their own. The environment of two children living in the same country under anything but exceptional conditions probably resembles a great deal more than it differs. The mere fact of a 12-year-old child's ignorance of the names of the month is not, *per se*, a reflection upon the child's intelligence. The child may be deaf, or he may have been reared by feeble-minded parents, or he may never have attended a school. But if the 12-year-old has had the usual home and school advantages

common to a civilized people such a deficiency as that noted becomes decidedly significant. Even then, however, we should not be justified in pronouncing the child subnormal upon this evidence alone. Gaps may occur in the education of anyone. It is only when the child's intelligence has been triangulated from many positions, and always to its disadvantage, that we can reasonably infer the presence of a subnormality marking the subject as incapable of acquiring the common elements of our social inheritance.

Surveys of this nature are possible only in so far as we have age norms for different lines of performance, few of which have yet been established. Many studies have been made showing age and grade tendencies for certain kinds of mental ability, but for the most part these have either been too limited in scope, or have not utilized non-selected subjects, or else have been made under conditions impossible to duplicate. Worst of all, the experimentation has been of haphazard nature instead of being guided by some directing idea like that of establishing definite age or grade standards. The Binet tests of 1908 are the only set hitherto devised covering any considerable variety of functions or directed by any comprehensive plan. The invaluable contribution of Binet consisted in the simple device of arranging the tests in a series of groups according to their difficulty as determined by age differences in performance. In spite of the limited data which served as the basis for his scale, the 1908 tests of intelligence immediately became serviceable and already seem to have proved their worth to all who have experimented with them. The series can readily, by extensive and thorough trial, be indefinitely extended, refined and adapted to different conditions and purposes. The following preliminary and imperfect study is offered as a slight contribution to this end.[2]

The purposes of this study are (1) to determine the adaptability of the Binet tests to American conditions; (2) to discover the changes in serial arrangement that may be necessary

[2]This study was completed before the publication of Binet's 1911 revision, and before the appearance of Goddard's and Bobertag's criticisms.—AUTHORS.

to make them reliable; (3) to try out certain other tests designed to supplement the Binet series, and (4) to arrive at some conclusion as to the psychological and pedagogical value of the tests individually and collectively.

Sources of Data and Conditions of the Study.

Between February 1 and May 1, 1911, the writers, with the assistance of Miss Katherine Kip and Miss Edith M. Bushnell, gave the Binet tests to 396 non-selected children.[3] The schools selected for the tests were a district school on the Stanford University campus, attended almost equally by the children of college professors and of laborers (30 children); the public schools of the village of Mayfield, Cal., where about one-half of the entire number were tested (72); the city schools of Palo Alto (150); a rural school near Stanford University (20), and two public kindergartens at Long Beach, Cal. (124). In the campus school, rural school and kindergartens practically all the pupils were tested; in Palo Alto the tests were given to the entire third and fourth grades of one building, to one-half the A and B fifth grades, and to several in each of the first, second, sixth and seventh grades. Where not all the pupils of a room were tested, as was the case in the Mayfield and some of the Palo Alto work, care was taken to secure representative case by requesting teachers to send only pupils whom they considered of average intelligence, or, in case a pupil of either exceptionally high or low ability was submitted, to follow this by another varying in the opposite direction from the normal. It is believed that the material is about as representative, with the exceptions later to be designated, as could be secured. As far as home advantages and social milieu are concerned the pupils are probably not very different, on the whole, from what would be found in the better class villages

[3]The tests were all made under the supervision of Dr. Terman, who himself applied the tests to about 90 children. Two hundred and twenty-two pupils were tested by Mr. Childs, and 124 by Miss Kip and Miss Bushnell. About 40 of the cases tested by Dr. Terman were a selected retarded group, the data from which could not be included in this study. The writers are greatly indebted to Miss Kip and Miss Bushnell for their very able and exceedingly careful assistance.

and towns from New England to California. The age and grade distribution was as follows:—

Age.	Number.	Grade Distribution.
4 years	29	27 (Kgn), 2 not in school.
5 years	83	81 (Kgn), 2 not in school.
6 years	26	17 (Kgn), 9 (I).
7 years	29	13 (I), 12 (II), 4 (III).
8 years	43	7 (I), 15 (II), 16 (III), 5 (IV).
9 years	49	2 (I), 2 (II), 16 (III), 26 (IV), 3 (V).
10 years	33	6 (III), 13 (IV), 11 (V), 3 (VI).
11 years	44	1 (II), 3 (III), 13 (IV), 18 (V), 7 (VI), 2 (VII).
12 years	35	6 (IV), 11 (V), 2 (VI), 7 (VII).
13 years	17	1 (IV), 5 (V), 4 (VI), 6 (VII), 1 (VIII).
14 years	6	3 (VI), 3 (VIII).
15 years[1]	2	1 (VI), 1 (VIII).

[1]Age here means age at last birthday; *i. e.*, 4 years includes children between 4 and 5, etc.

The tests were conducted in hallways or vacant rooms of the various school buildings, practically free from distractions. It was impossible to observe any uniformity in regard to time of day, the time ranging from 8.30 A. M. to 4.30 P. M. It is doubtful, however, whether this factor vitiates the results very materially, inasmuch as the tests almost never failed to enlist the interest of the child. The experiments of Thorndike and others indicate that undoubtedly existent fatigue does not materially influence a brief intellectual performance which holds the S's spontaneous attention. The time required for the tests was usually between 30 and 45 minutes, depending upon age, the number of tests covered and the promptness of responses.[4] Extreme care was taken to win the confidence of the child and to rid him of any embarrassment before beginning the tests. Occasionally several minutes were consumed in this way.

The kindergarten children took all the tests from the third to the eighth year, inclusive. The others, with a few exceptions, began the series a year below the S's chronological age, or further back if necessary, and continued as follows: The first and second grade pupils generally through the ninth-year

[4]The testing of seven to ten pupils constitutes a day's work of five hours if the examination is made reasonably deliberate and thorough. In the opinion of the authors, tests carried through at the rate of 20 to 30 per day are sure to give unreliable and misleading results.

group, and in many cases through the tenth year. In the grades above the second, testing usually began with the eighth-year group, but farther back if necessary, and proceeded as far as the pupil could go. In some of the earlier tests a few were not taken over as large a range as was later covered, and some of our data, therefore, may not be absolutely complete. Discarding some of the more questionable records has left the data here included, the writers believe, not seriously deficient from this cause.

Too much emphasis cannot be placed upon the importance of securing uniformity of conditions and procedure in applying the tests. An apparently insignificant change in the wording of a question or of a preliminary explanation may influence the response very appreciably. It is inevitable that, in spite of all precaution, a preliminary study of this kind will contain some errors arising from this source, considering that even variations in the vocal inflection of different E's will influence the result. Elaborate instructions, however, were placed in the hands of each person who proposed to assist in the tests, and these were carefully studied before the work began. Moreover, the methods of different workers were occasionally checked up by comparison during the progress of the tests.

The detailed instructions used are too bulky for reproduction here, but were practically identical with those in Whipple's Manual, except on the following points:

(1) The test of "making change" involved the purchase of a 9-cent article and payment with a 25-cent piece.

(2) For the test of ability to name pieces of money the following eight coins were used: Penny, nickel, dime, quarter, half-dollar, dollar, five dollars (gold), ten dollars (gold).

(3) *Palo Alto, river* and *money* were the three words used for sentence construction.

(4) The following tests of "comprehension" were used in year five:

(*a*) What's the thing to do when you feel sleepy?

(*b*) What's the thing to do when you feel cold?

(*c*) What's the thing to do if it's raining when you start to school? (Two out of three correct.)

The following were used in year ten:

(*a*) What ought you to do when you have missed a train?

(*b*) What ought you to do when you have been struck by a playmate who did not do it purposely?

(*c*) When you have broken something which does not belong to you?

(*d*) When you have been detained so that you will be late for school?

(*e*) Before taking part in an important affair?

(*f*) Why does one excuse a wrong act committed in anger more readily than a wrong act committed without anger?

(*g*) What ought you to do when asked your opinion about someone whom you only know a little?

(*h*) Why ought you to judge a person more by his acts than by his words?

(Five out of eight correct.)

(5) Numbers 1, 2, 4 and 5 of Whipple's absurdities (see Whipple's Manual, p. 509) were used, and in addition the following: "There was a railroad accident yesterday, but it was not serious. The number of dead is only 48."

(6) In the tests of memory for sentences and also for digits one

(7) In the rhyming test "day" was substituted for "obey." Year 12.)

(8) The following "problems of fact" were used in year 12:

1. "There was a little boy who had never been to the city. When he was six years old his father took him to San Francisco. As soon as the boy saw the electric street cars for the first time, he said . . ." "What do you think the boy said?"

2. "My neighbor has just been having queer visitors: first a doctor, then a lawyer, then a priest. What's happening there?"

3. "An Indian, coming to town for the first time, looked very intently at a white man riding up the street. When the white man had passed the Indian said to his companion: 'White man lazy; him walk sitting down.' What was the white man sitting on?"

(1 and 3 are substituted for the rather too gruesome problems of Binet.)

An S's test age was computed by assigning him to the lowest year in which he passed all the tests, or all but one, and then crediting him with a half year additional for each three tests satisfactorily passed beyond this. In the seventh year four tests were considered as the equivalent of a half year. From the ninth to the twelfth years, inclusive, five credits were counted as equivalent to a year, and a remainder of three as a half year. This seems to be the only way of giving the tests of any year approximately equal value with those of any other year if Binet's method is to be adhered to at all. Even by this method the child often loses credit for one or two tests passed beyond the half year to which he has been assigned, but on the

other hand he often gains one credit where he fails in but one test of a given year. Binet's method of assigning credit is crude at the best. The changes here introduced below the ninth year tend rather to draw the results nearer to the Binet norm than away from it. A much more accurate method will be suggested in the last article of this series. ,

Another difference between the method here employed in reckoning test age and that followed by Binet and others touches a more fundamental point and deserves special consideration. In deriving the age norms for the various tests it seems that Binet reckoned chronological age by even years; that is, by 8-year-olds, he means all children between 8 and 9 years; by 12-year-olds, all between 12 and 13, etc. In other words, the median age for the so-called 8-year group is about 8½ years. This is the most common procedure, is the one used also in our tests, and is perfectly legitimate, provided the facts are kept in mind. This means that what Binet designates as a characteristic performance of 8 years is really that of 8½ years. It follows, therefore, that when an S (whatever his chronological age) passes as far as to the end of the 8-year group, and no farther, or if his performance is equivalent to this, then such S should be assigned a test age of 8½ years. This was the method of computing used in this study, one which obviously throws our test ages one-half year in advance of those which would have been assigned by the method others have used. *This should be held in mind when comparisons are made of our results and those obtained by other studies.* The only logical way to avoid the necessity of adding on this half year is to group differently the S's to whom tests are applied for the sake of securing age norms. They can be grouped either by half years or in year groups separated by the half-year point. If we define the standard for 8-year mentality as the characteristic performance of non-selected S's lying between 7½ and 8½ years, then we can continue to use the Binet method of reckoning. Otherwise it is necessary to rename his standards as those for 6½, 7½, 8½, 9½ years, etc., instead of calling them, as Binet does, norms for 6, 7, 8, 9 years, etc. To illustrate the working of this, an S completing all the tests to the middle of the 8-year group would be ranked as testing 8

years. If he passes three additional tests either in this or the following years, he would be ranked as testing 8½ years.

Results from Tests of 396 Pupils.

Although, generally speaking, it is desirable to publish all the raw data from investigations of this nature, it is doubtful whether a preliminary excursion of this kind into a field where conditions and procedure have not been fully standardized warrants the expense and trouble of publishing the twenty or more pages which would be required for our material. We reserve this, together with a more thorough mathematical treatment of the data, for a later and more extended study.

The following table shows the distribution of *test ages* resulting from the application of the scale to the 396 subjects:

TABLE II.

Age.	No. tested.	Variations of test age from chronological age.													Median age.	Median test age.
		+3.0	+2.5	+2.0	+1.5	+1.0	+0.5	0.0	−0.5	−1.0	−1.5	−2.0	−2.5	−3.0	−3.5	
4	29		8	13	4	3	1									4.75	6.5
5	83		3	14	31	24	5	5	1							5.5	7.0
6	26				2	9	7	7		1						6.37	7.0
7	29		1	1	6	6	4	8	2		1					7.5	8.0
8	43	2	1	3	2	6	6	7	6	7	3					8.5	8.5
9	49			5	6	5	6	13	8	3	3					9.5	9.5
10	33			1	3	5	5	7	5	3	4					10.5	10.5
11	44					1	3	10	7	11	5	5		2		11.46	10.5
12	35							3	2	10	7	7	3	2	1	12.33	11.0
13	17								1	2	1	8	1	3	1	13.42	11.5
14	6										1	3		2		14.58	12.5
15	2														2	15.2	12.0

Our data, so far as the Binet tests are concerned, are not at all conclusive for the thirteenth year, since we did not get the brightest pupils of this age for the reason that our tests were not extended into the eighth grade. The number of cases at fourteen and fifteen years is altogether too small to warrant any conclusions relative to these ages.

The following table shows a brief comparison of results by years and by grades:

TABLE III.

Age.	Median age.	Median test age.	School grade.	Median grade age.	Median grade test age.
7	7.5	8.0	I	7.5	7.5
8	8.5	8.5	II	8.17	8.3
9	9.5	9.5	III	9.33	9.5
10	10.5	10.5	IV	10.0	10.0
11	11.46	10.5	V	11.42	10.7
12	12.33	11.0	VI	12.17	11.5
*13	13.42	11.5	†VII	12.5	12.5

*Based on rather limited data, 17 cases.
†Based on rather limited data, 15 cases.

The following chart represents graphically the relation between the median chronolgical and the median test ages, and throws into relief the periods of rapid, normal and slow mental growths as measured by the Binet scale. The points in the upper horizontal line represent chronological and median ages, and in the lower line test ages and median test ages. The connecting lines join the corresponding medians. When there is an identity in the two medians the connecting line is vertical. If an age group tests ahead of the chronological age, the connecting line slants to the right as it goes down; if the group tests behind the age, to the left. It will be noted that the chronological ages cover a range of nine years; the test ages but five years. In other words, the scale is far too easy at the lower end, while at the upper end it is too difficult. An S at the upper range is also at a great disadvantage in that he is not given opportunity to try tests beyond thirteen years.

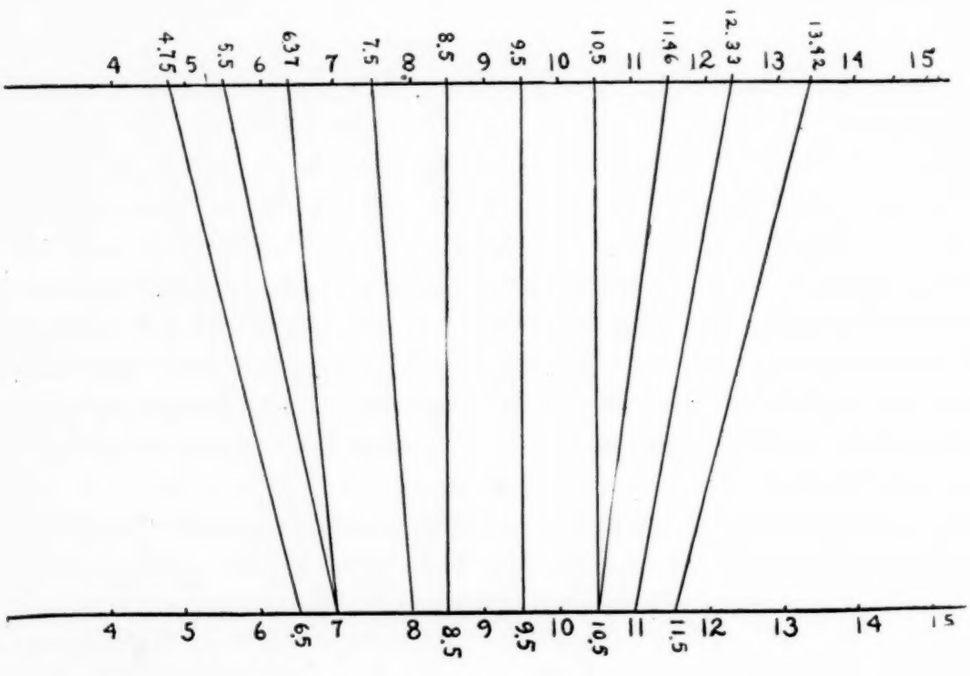

FIGURE I.—Upper line, chronological ages; lower line, test ages.

Charting the test ages as Dr. H. H. Goddard has done, by lumping all the ages together (see *The Training School* for

May, 1911), conceals, of course, the very facts we wish to know. From the above it is seen that the number of younger pupils testing ahead is about balanced by the number of older ones testing behind. What we want to know is how nearly accurate the scale is at every point. The following chart shows graphically how badly the scale fits California children of five and twelve years:

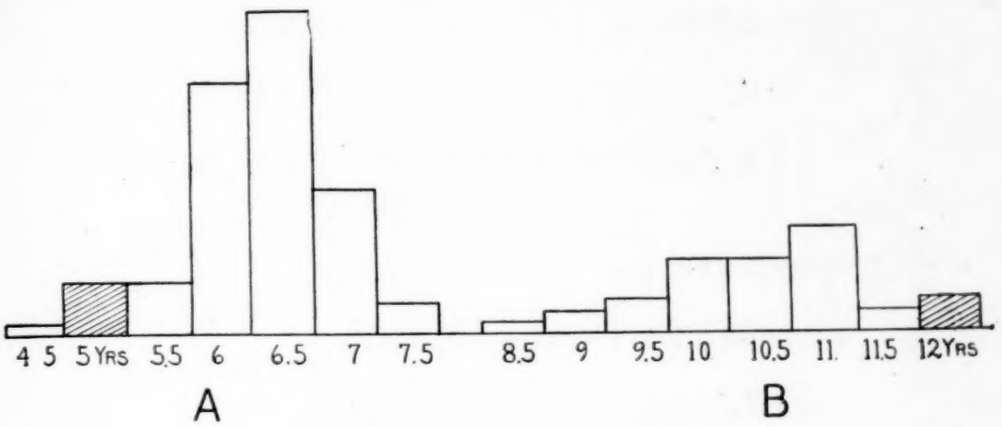

FIGURE II.—"A" shows the distribution of "test ages" for 83 five-year-old children; "B," for 35 twelve-year-olds. In each case the number testing normal is indicated by the barred column.

Table IV shows the number of children tested at each age and the per cent. of those tested at each age who pass each test taken. A comparison with Binet's scale will show how erroneously some of these tests are placed for use with American children. Our results indicate that for California children, at least, some of the tests will have to be moved as much as three years.

In the table where at any given year the number tested is but a small part of the whole number of that age the results should not be given serious consideration, for these represent either retarded older pupils who were given the lower year tests or the precocious younger children who were carried to a point far in advance of the majority of their age.

Further criticisms and suggestions are reserved for a later article.

TABLE IV.

This table shows the per cent. of those trying a test at any particular age who passed it. The figures in heavy type show the number of pupils of each age who tried each test.

IV.

Age.	No. of cases.	1	2	3	4	5
4	**29**	100	93	97	100	93
5	**83**	100	100	100	100	97
6	**18**	100	100	100	100	83
7	**0**	...	...	...	...	...
8	**0**	...	...	...	...	...
9	**0**	...	...	...	...	...
10	**0**	...	...	...	...	...
11	**0**	...	...	...	...	...
12	**0**	...	...	...	...	...
13	**0**	...	...	...	...	...
14	**0**	...	...	...	...	...
15	**0**	...	...	...	...	...

V.

No. of cases.	1	2	3	4	5	6
29	97	86	90	90	51	97
83	98	91	88	91	83	96
20	95	85	95	90	85	95
4	100	100	75	100	75	100
2	0	100	50	100	100	100
0	...	...	...	...	...	...
0	...	...	...	...	...	...
0	...	...	...	...	...	...
0	...	...	...	...	...	...
0	...	...	...	...	...	...
0	...	...	...	...	...	...
0	...	...	...	...	...	...

VI.

No. of cases.	1	2	3	4	5	6	7
29	41	80	31	86	83	72	65
83	73	88	53	91	81	79	84
24	75	83	62	83	87	96	87
20	90	90	80	90	95	95	100
14	93	93	72	93	100	100	100
3	100	100	100	100	100	100	100
0	...	...	...	...	...	...	...
0	...	...	...	...	...	...	...
0	...	...	...	...	...	...	...
0	...	...	...	...	...	...	...
0	...	...	...	...	...	...	...
0	...	...	...	...	...	...	...

VIII.

No. of cases.	1	2	3	4	5	6
29	0	0	65	0	0	34
88	0	3	74	1	0	47
26	0	11	84	7	7	47
29	20	58	96	62	79	55
43	51	60	100	69	83	65
45	89	73	100	95	100	89
20	90	85	100	100	100	100
14	86	72	100	93	100	100
6	100	83	100	100	100	100
0	...	...	...	...	...	...
0	...	...	...	...	...	...
0	...	...	...	...	...	...

VII.

No. of cases.	1	2	3	4	5	6	7	8
29	34	14	0	23	17	44	34	17
83	68	33	10	53	50	47	55	38
26	57	73	42	61	50	42	77	53
25	68	92	100	92	72	60	100	72
35	82	97	100	91	74	60	97	88
12	83	100	100	92	83	83	100	92
1	100	100	100	100	0	100	100	100
2	100	100	100	100	100	100	100	100
0	...	...	...	...	...	...	...	...
0	...	...	...	...	...	...	...	...
0	...	...	...	...	...	...	...	...
0	...	...	...	...	...	...	...	...

TABLE IV.—Continued.

This table shows the per cent. of those trying a test at any particular age who passed it. The figures in heavy type show the number of pupils of each age who tried each test.

IX.

No. of cases.	1	2	3	4	5	6
0	...	...	...	...	...	...
0	...	...	...	...	...	...
0	...	...	...	...	...	...
27	25	85	7	3	0	25
38	52	85	26	29	13	29
49	81	98	42	24	46	55
31	77	100	35	58	58	54
39	87	100	67	61	65	77
29	93	100	62	69	62	55
12	92	100	83	83	100	75
2	100	100	100	100	100	50
0	...	...	...	...	...	...

X.

No. of cases.	1	2	3	4	5
0	...	...	...	...	...
0	...	...	...	...	...
0	...	...	...	...	...
16	62	18	18	6	31
28	64	50	42	17	50
46	91	69	60	37	65
33	81	88	84	57	92
43	81	96	86	64	71
35	93	87	93	77	78
17	94	94	100	89	89
6	100	100	100	83	100
2	100	100	100	100	100

XI.

No. of cases.	1	2	3	4
0	...	...	...	...
0	...	...	...	...
0	...	...	...	...
4	25	50	0	0
20	50	35	10	10
42	40	57	2	19
33	52	67	18	24
43	55	83	16	32
34	73	82	26	62
17	88	94	36	66
6	100	67	33	83
2	50	100	100	50

XII.

No. of cases.	1	2	3	4
0	...	...	...	...
0	...	...	...	...
0	...	...	...	...
3	0	100	0	0
14	28	85	42	7
27	29	74	29	18
28	46	92	35	35
37	43	81	37	40
34	38	82	41	41
17	36	89	54	48
6	83	100	67	100
2	50	50	100	100

XIII.

No. of cases.	1	2	3
0	...	...	...
0	...	...	...
0	...	...	...
0	...	...	...
0	...	...	...
0	...	...	...
0	...	...	...
0	...	...	...
8	25	...	25
19	15	...	30
9	22	...	55
5	40	...	60
2	0	...	50

EXPLANATION OF TABLE IV.

The Roman headings in the above table designate year group; the Arabic designate the individual tests in the following order:

IV—1. Sex. 2. Names familiar objects. 3. Three digits. 4. Comparison of two lines. 5. Sentences of ten syllables.

V—1. Weights. 2. Copies square. 3. Divided rectangle. 4. Counts four. 5. Fourteen syllables. 6. Comprehension questions.

VI—1. Right hand, left ear. 2. Aesthetic comparison. 3. Sixteen syllables. 4. Defintions of familiar objects. 5. Three commands. 6. Knows age. 7. Morning and afternoon.

VII—1. Omissions from pictures. 2. Number of fingers. 3. Writing from copy. 4. Diamond. 5. Five digits. 6. Description of pictures. 7. Thirteen pennies. 8. Four coins.

VIII—1. Reading for two memories. 2. Value of stamps. 3. Four colors. 4. Counting 20 to 0. 5. Writing from dictation. 6. Comparison of things.

IX—1. Names date. 2. Days of week. 3. Makes change. 4. Definitions superior to use. 5. Six memories. 6. Weights.

X—1. Names months. 2. Nine pieces of money. 3. Three words in sentence. 4. Comprehension questions. 5. Six digits.

XI—1. Absurd statements. 2. Sixty words in two minutes. 3. Defines abstract terms. 4. Disarranged sentences.

XII—1. Seven digits. 2. Rhymes. 3. Twenty-six syllables. 4. Problems of fact.

XIII—1. Draws from design in folded paper. 2. Reversed triangle. 3. Distinctions between words.

MENTAL FATIGUE IN DAY SCHOOL CHILDREN AS MEASURED BY IMMEDIATE MEMORY. PART II.

W. H. WINCH.
London, England.

SUMMARY.

Tests of immediate memory on a class of fifty-one boys, averaging eleven years of age, showed a slight relative inefficiency in the work of the late afternoon as compared with that of the early morning. The decrease in efficiency was about two per cent.

IV. A Second Experiment in a Boys' School.

This experiment was carried out with the whole of a Standard IV class, numbering 51, of an average age of 11 years, 1 month, on June 30, 1911. The school was the same as that in which the previous experiment was done, but the standard of mental proficiency and the age of the boys were different and they were taught by a different teacher. This class was also taught by a young man of capacity, but differing from the teacher of the preceding class somewhat in the greater evenness, though, perhaps, slighter intensity of his demands upon the mental energy of his pupils.

i. *Tests and Method of Marking.*

The tests were tests of memory of auditory percepts. Six consonants, for example, r, l, p, x, m, s, were called out twice in 25 seconds by the teacher of the class, whose voice and enunciation were familiar to the children. A slight pause was made after each consonant with a slightly longer pause after the third. One minute five seconds was allowed for the boys to write out as much as they could remember of each set of con-

sonants immediately after they had been called out. They were required to write them in two lines of three consonants each. Ten sets of consonants constituted one test or exercise, the whole procedure lasting exactly 15 minutes. The boys' papers of answers were marked exactly as in the preceding experiment. Three marks were given for each consonant correctly remembered and rightly placed; two marks, if the consonant were rightly remembered, but was one place before or behind its proper position; and one, if it were two places out. As before, the children were instructed to leave a space in the place of any consonant which they knew they had forgotton.

ii. *Chronology of the Series.*

Four sets of preliminary tests were given to afford a basis for the division of the class into two equal groups. They were worked from 10.45 to 11 A. M. on Tuesday, June 13; Wednesday, June 14; Friday, June 16, and Tuesday, June 20.

The school lessons which preceded these preliminary tests were in all cases the same, namely, Scripture, arithmetic and recreation, so that no difficulty arises on that score, but I regret that the exigencies of the school work in relation to the forthcoming coronation festivities prevented me from spreading the work over a longer period and, above all, of getting more regular intervals between the successive tests. Coronation holidays followed immediately upon the preliminary tests, and it was not until some twelve days afterwards that the school once more settled down to work. In the interim the boys taking the experiment had been divided into two equal groups on the results of the preliminary tests and henceforward one of the groups continued the tests in the mornings from 9.45 to 10 A. M., whilst the other group worked the same tests in the afternoons of the same days from 4 to 4.15 P. M.

Four of these tests were given—henceforward called the final tests—on Wednesday, July 5; Friday, July 7; Wednesday, July 12, and Friday, July 14.

The lessons which preceded the morning tests were in each case Scripture lessons: the lessons which preceded the afternoon tests were, on Wednesdays, arithmetic, physical exer-

cises and reading, and on Fridays, English, physical exercises and history.

Whilst Group A were working their tests in the morning, Group B, in another room, were drawing common objects, and whilst Group B worked their memory tests in the afternoon Group A drew common objects. The children in both groups were told the results of their previous work both in drawing and in memory work before proceeding to the next test.

In every respect other than those just indicated the time-table of school work for the two groups was precisely the same throughout the period of the experiment.

iii. *Results.*

It is of little value to attempt to divide a class into equal groups on results which vary so much from exercise to exercise that the children seem to be jumping about in the lists and rapidly exchanging places. An inspection of the results will give a fairly good notion as to the "steadiness" of the work, but, of course, a calculation of a coefficient of correlation enables more careful comparison to be made. Calculated on the moment-product formula $r = \frac{\Sigma xy}{N\sigma_1\sigma_2}$, the correlation coefficient between the results of tests 1 and 2 is + .743, of tests 2 and 3 is + .741 and of tests 3 and 4 is + .746. With correlations of this size and a number of cases amounting to 51, the "probable error" of the coefficient of correlation is practically negligible. I decided to make the division on the results of these four tests, but I was well aware that I had not divided the class in the case of the previous experiment until the positive correlations between the results of successive tests had become decidedly higher than these, and I should have continued the preliminary tests had not the coronation holidays stood in the way. It is probable, therefore, that the division into equal groups is not so satisfactorily effected as in the case of the previous experiment. That it is fairly satisfactory, however, will appear from tables which follow.

I propose first to show the work of the two groups as wholes both in the preliminary and final tests.

*TABLE I, *Showing the Work of the Two Groups as Wholes, Test by Test, in Both the Preliminary and Final Tests.*

		Preliminary Tests.				Final Tests.			
		1st.	2d.	3d.	4th.	1st.	2d.	3d.	4th.
Group A.	Av. mark....	15.4	16.1	16.0	16.0	16.6	16.2	16.4	16.9
	M. V.....	1.4	1.2	1.5	1.4	1.2	1.6	1.2	1.1
Group B.	Av. mark....	15.2	16.2	16.0	16.0	16.0	16.1	16.1	16.3
	M. V.....	1.4	1.3	1.6	1.2	1.4	1.1	1.2	1.3

The average mark for the whole of the preliminary tests for Group A is 15.9, with a mean variation of 1.2, and for Group B is also 15.9, with a mean variation of 1.2. In the final tests Group A scores an average mark of 16.5 (mean variation 1.1) and Group B an average mark of 16.1 (mean variation 1.1). Group A—the morning group—improves on its preliminary record to the extent of 3 per cent.: Group B—the afternoon group—improves on its preliminary record to the extent of 1 per cent. There is apparently an advantage of 2 per cent. in improvability of the morning over the afternoon workers.

There are one or two comments which may usefully be made on the above table. It will be noticed how little improvement is shown in the preliminary tests, even though the exercises are quite new. But surprise will no longer be felt when it is remembered that they were all worked within the space of seven days. An improvement in mental work of this kind is very little evident unless there are suitable intervals between the successive exercises. It will also be seen that the morning group did better work in each of the final tests as compared with the group of afternoon workers, though, in one case, the difference is slight.

I next propose to show, section by section, the work of the two groups in the preliminary and final tests respectively.

*The calculations are made to the nearest ten.

*TABLE II, *Showing the Work of Groups A and B Compared, Section by Section, in the Work of the Preliminary and Final Tests, Respectively.*

Marks for four preliminary tests.	Group A. No. of boys.	Group A. Av. mark per boy per test of four prelim'y tests.	Group A. Av. mark per boy per test of four final tests.	Group B. No. of boys.	Group B. Av. mark per boy per test of four prelim'y tests.	Group B. Av. mark per boy per test of four final tests.
Over 65.....	12	17.1	17.5	12	17.1	17.0
Over 60.....	7	15.8	16.7	7	15.7	16.1
Over 55.....	2	14.0	15.5	2	14.6	15.1
Under 55.....	4	13.3	13.8	4	13.1	14.1

There appears to be some advantage on the side of the early morning workers in the first, second, and third sections; but in the lowest sections, consisting of boys who obtain less than 55 marks in the four preliminary tests, the group of afternoon workers show more improvement than those who work early in the mornng. This is contrary to the suggestion, previously made, that, perhaps, it is the more strongly endowed who do better work in the afternoon, for these boys are among the weakest in the class mnemonically, as shown by both the preliminary and final tests. The differences between the improvements shown by the morning and afternoon workers appear small. This may be partly due to the fact that the tests were a little too easy. Eighteen (or rather one hundred and eighty) is the maximum mark per test and it will be at once seen that in the highest sections, which score 17.1 marks in the preliminary tests, that there was not very much room for improvement. I propose, finally, to show the comparative improvements of the individual pupils of Group A and Group B. The improvements are calculated on the average of the four preliminary tests on which the class was divided. The averages and percentages are calculated to the nearest unit only.

*The calculations are made to the nearest ten.

TABLE III, *Showing the Comparative Improvement of the Pupils in Groups A and B in the Morning and Afternoon Work, Respectively.*

Group A.				Group B.			
Name. (Initials only.)	Av. of four prelim'y tests.	Av. of four final tests.	Percentage of improvement.	Name. (Initials only.)	Av. of four prelim'y tests.	Av. of four final tests.	Percentage of improvement.
I. F.....	177	176	— 1	S. S.....	177	180	+ 2
F. C.....	172	179	+ 4	W. V.....	175	174	— 1
F. G.....	171	174	+ 2	M. A.....	174	174	0
R. D.....	173	173	0	G. H.....	172	170	— 1
G. E.....	172	176	+ 2	O. S.....	172	178	+ 3
P. A.....	170	170	0	W. W.....	172	177	+ 3
P. G.....	168	173	+ 3	R. R.....	170	169	— 1
F. W.....	170	176	+ 4	P. W.....	168	175	+ 4
F. A.....	164	176	+ 7	N. L.....	165	162	— 2
H. W.....	168	173	+ 3	S. W.....	165	159	— 4
R. H.....	167	175	+ 5	O. S.....	166	154	— 7
K. J.....	164	166	+ 1	McW.	164	171	+ 4
S. A.....	162	179	+ 10	S. T.....	161	161	0
H. J.....	160	168	+ 5	P. A.....	158	155	— 2
D. A.....	157	163	+ 4	H. W.....	160	169	+ 6
C. L.....	155	165	+ 6	D. R.....	159	172	+ 8
A. J.....	158	167	+ 6	C. J.....	153	163	+ 7
S. K.....	155	158	+ 2	M. A.....	153	144	— 6
E. W.....	154	162	+ 5	N. W.....	151	160	+ 6
N. W.....	138	144	+ 4	McI. F.....	148	152	+ 3
W. E.....	142	164	+ 15	H. C.....	143	151	+ 6
W. E.....	136	147	+ 8	B. T.....	136	136	0
B. S.....	133	148	+ 11	D. H.....	131	154	+ 18
W. F.....	132	126	— 5	H. E.....	130	136	+ 5
H. A.....	127	128	+ 1	R. F.....	123	132	+ 7
Totals...3945		4106		Totals...3946		4028	
Averages. 157.8		162.2		Averages. 157.8		161.1	
M. V..... 11.9		11.1		M. V..... 12.2		11.2	

An inspection of the table shows that in the group of morning workers improvement is shown in every case but four, whilst, in the afternoon group there are 11 such cases; but probably comparisons of this kind may be made more readily from the following tabular statement.

TABLE IV, *Showing the Comparative Percentages of Improvement Between the Members of Group A and Those of Group B.*

	Number of Cases.	
	Group A. Morning Group.	Group B. Afternoon Group.
Gain of 15% and over	1	1
Gain of 10% to 15%	2	0
Gain of 5% to 10%	7	7
Gain of 0% to 5%	11	6
Neither gain nor loss	2	3
Loss of 0% to 5%	2	6
Loss of 5% to 10%	0	2

The coefficient of correlation on the total results of the four preliminary tests between the corresponding cases (see Table III), of the two Groups A and B is, of course, as estimated by rank formulæ + 1. On the moment-product formula, which uses marks, not rankings, the coefficient is + .899. Similarly calculated, the coefficient for the final tests is + .862.

V. SUMMARIZED CONCLUSIONS.

1. Again we find that the use of immediate memory tests indicates, in a normal group of school children constituting the whole of a class or grade, a certain amount of relative inefficiency in late afternoon as compared with early morning work.

If we accept this relative inefficiency as a measure of fatigue for mechanical memorizing it appears to be very small for children of this age and mental proficiency, namely, about 2 per cent. Unfortnately, however, the tests set were relatively, as well as absolutely, easier than those given to the older and abler boys who took the previous experiment, so that I feel very little confidence in making the suggestion that with lower classes the fatigue effects for purely mechanical work may be smaller than in higher ones. In any case the fatigue effects for mechanical mental work of this kind appear small amongst children, at least so far as boys of this type are concerned.

2. As in the class previously dealt with there are to be found children who cease to improve and tend downward. It is not argued that no fatigue is present if improvement continues, indeed, I have tried to measure fatigue by differential improvements made under differing conditions; but it is ar-

gued that an actual deterioration in work of this kind shows such fatigue that there are indications to the educationist of the advisability of a cessation for a time of that or other work or what amounts to much the same thing, of a longer interval between the successive exercises. Two of the morning workers and eight of the afternoon workers appear to have deteriorated during the course of the work. With one-third of the afternoon group, therefore, the work appears to have been futile, as was found to be the case in the previous experiment. On the other hand a considerable proportion of the children have continued to improve.

3. The ordinarily accepted pedagogical maxim that, with most children, mental work of a mechanical kind may be done, with fair satisfaction, late in an afternoon session, appears to receive some corroboration.

ADDING UPWARD AND DOWNWARD.[1]

LAWRENCE W. COLE,
Professor of Psychology and Education, University of Colorado.
University of Colorado.

SUMMARY.

Measured by averages, twenty-nine out of a group of thirty[1a] persons, selected at random, added the same problems more rapidly and less accurately when adding upward than when adding downward. Counting to the left was also slower and more accurate than counting to the right. This is apparently due to a habit acquired by reading from left to right. In both experiments the factor of habit seemed to produce a saving of time at the expense of accuracy. The subjects of the experiments were persons of average practice in adding. Probably a very great amount of drill in addition (perhaps more than a school program could allot to it) would be required before results of the opposite type would emerge and the adding process become both mechanical and accurate.

INTRODUCTION.

A statement by Professor Muensterberg[2] concerning "the considerable time differences between even the simplest subtraction and addition processes" led me to think it possible that school training in arithmetic may give to a majority of pupils a habit of adding *upward* which would make it impossible for them to add the same columns as rapidly *downward*. Should this prove to be the case, it seemed possible to ascertain what proportion of the effect of practice is due to central processes and what part due to peripheral ones. Unless my ran-

[1]My thanks are due to my assistants, Mr. C. I. DeVoss and Miss Mary Lakenan for valuable help in the experiments.

[1a]Five other persons took the preliminary test; hence I mention 35 below.

[2]Psychology and the Teacher, p. 285.

dom selection of thirty-five persons is an exceptional one, or the number of persons too few, the experiments seem to show that such a habit often exists, and this appears from the record of the time averages, of accuracy, of persons, and from the decrease of reversals of the habitual time relation with an increase of adding time. Until after experiments began I did not expect to find a greater difference in accuracy than in time between upward and downward addition.

Method.

Preliminary Test: In a preliminary test five persons, four advanced students of psychology and the writer, each added four five-column problems and four six-column problems, twenty figures in length, both upward and downward. The order of the additions was such as to give no advantage to either the upward or the downward process. Defects in this test led to the arrangement of Tests A and B for further work.

Test A: This consisted of twenty single columns, forty figures in length, so arranged that those designated by the odd numbers, 1-19, inclusive, were to be added upward, and those designated by the even numbers, 2-20, inclusive, were to be added downward, but the subject added the columns in consecutive order from 1 to 20. By this arrangement columns 1 and 12, 2 and 11, 3 and 14, etc., which were the same, were added in opposite directions. So not fewer than eight nor more than eleven columns intervened between any two to be added in opposite directions. Of those to be added in opposite directions half were first added downward and half were first added upward.

On the same page with these typewritten columns there was a preliminary exercise in three-column addition, twenty figures in length, to be added, first upward, then downward. This was done in order to be sure that the subject understood the directions given him, and in the hope that any initial flurry or embarrassment would be overcome during this exercise. The times required to get these two sums were recorded, but used no further.

Test B: In Test A the ten columns to be added upward were

identical with the ten to be added downward. Hence, in an upward addition, the subject would add 2, 6, 3, etc., and in the corresponding downward one 3, 6, 2, etc. As measured by time, accuracy and "reversals" (and somewhat vaguely by introspection) these two processes are not quite the same. Consequently, in Test B the columns to be added downward *were inverted.* So in both the upward addition and the corresponding downward one the subject added the successive figures *in the same order.* In other respects Tests A and B were the same.

Since the combinations to be made in the upward and downward columns of Test B were the same, and made in the same order, the only objective difference between the two types of addition was a difference *in the direction of the movements of the eye and hand.* Aside from fluctuations of attention, this difference would seem to be the cause of all differences in the results of the adding.

Test C: This experiment was not designed to test the effect of an addition habit, but of the habit acquired by reading from left to right. It consisted of two series, each of one hundred lines (more exactly the line used in typewriting to separate the numerator and denominator of a fraction; it is much like a plain figure 1). The first series was to be counted "by ones" from left to right; the second in the opposite direction. This, like Test B, involved objectively only a reversal of the habitual movements of hand and eye. Despite the small number of lines to be counted, the differences, both in time and accuracy of counting the two series, were relatively great.

Test C was doubtless defective in that it might suggest to some subjects that the lines to be counted leftward were the same in number as those which had just been counted to the right.

Method: In every case the experiment was made in a quiet room and the times carefully taken with a stop-watch. The subject was not urged to add "as quickly as he could," but he was directed to add as rapidly as possible without a sacrifice of accuracy. This, with the knowledge that his time was being recorded, served to arouse a strong effort to add both rapidly and accurately.

The same thirty persons took all three tests. Tests A and B were not given to the same person on the same day. The majority of the subjects were sophomores in college. Three or four were juniors and seniors. Twelve were men, eighteen were women.

Results.

Preliminary Test: The result of this experiment, in which five persons participated, showed that adding the six-column problems *downward* required an average of 15.3 per cent. more time than adding the same columns upward, while the downward addition of the five-column problems required 10.5 per cent. more time than their addition upward. Each average is obtained from the addition of twenty problems.

In the total of twenty pairs of problems there occurred only three "reversals," or cases of more rapid downward than upward addition. The records of these reversals are included, however, in the averages, which they reduce but slightly. One reversal was by a margin of 4 seconds, another of 8, and the third of 60 seconds. *The amount of error was less for adding downward.*

This test included two types of association, namely, "carrying" and adding. An error in the tens column had ten times as great a value as the same error in the units column, and carrying prevented reducing such errors to a common basis. As a final objection to the test, it required too much time to be given to a considerable number of persons. Aside from these defects, it served to differentiate the results of upward and downward addition very clearly. Nevertheless, I felt compelled both to shorten the time and to eliminate "carrying." After some further investigation I decided to depend on the difference of time between the upward and downward addition of ten single columns, each forty figures in length. A "reversal" in this case would mean that more time was required to add ten columns upward than downward. To fulfil these conditions Tests A and B were arranged. Each of them demanded about twenty minutes of adding by each subject, with a brief interval between each two consecutive columns.

Fortunately, in order to determine the suitableness of this amount of time, I can compare its effect on reversals with both a briefer and a longer period. This will at the same time give the results of the experiments as shown by reversals.

Reversals: When the time required to add a single column in both directions (the average time for this was 2 minutes, strictly 119.4 seconds) was used, reversals of time occurred in one-third of the pairs of columns. When measured by the time required to add *ten* columns in each direction (which averaged 20 minutes, or strictly 1194 seconds), reversals occurred in one-sixth (strictly 9/60) of the cases. If the addition was continued for an average time of 24 minutes and 22 seconds for each person, as in the five- and six-column addition (complicated by carrying), the reversals amounted to only one-seventh of the total number of pairs of problems added. It appears evident from this that in a group of persons who have formed the habit of adding more rapidly upward than downward the number of reversals, or cases of adding more rapidly downward than upward, can be made as few as we please by lengthening the columns to be added. My selection of ten columns, or about twenty minutes of adding time, therefore seems fair, though thirty minutes would have shown the effect of habit much more clearly.

The facts stated above may be tabulated as follows:

Unit.	Ratio of pairs of additions reversed to all pairs added.	Adding time.
Pairs of single columns	1/3	2 minutes.
Ten columns in each direction	1/6	20 minutes.
Five- and six-column addition of the preliminary test	1/7	25 minutes (nearly).

In one of the reversals of the twenty-minute period the subject was disturbed by hearing the signal for his next class. In another the subject felt sure her reversal was due to a headache produced by having arisen at 5 A. M. to play tennis. Like several others, these two persons reversed in Test B, but not in A, which they had taken on a previous day.

These facts seem to indicate that many of the reversals, and perhaps all but two, which were due to habit,[3] were caused by

fluctuations of attention. These, being momentary, would have far greater relative effect in very brief adding times. There seems to be no doubt that the shorter time for addition upward, which occurred in five-sixths of the pairs of twenty-minute additions, was due to a habit acquired when learning to add or when practicing for efficiency. Its influence on most of the pairs, whether measured by reversals, by time averages or by accuracy, would tend to show that the habit is not a momentary or intermittent cause, but a fairly constant one.

Record by Persons: In all, four persons gave reversals in Test A, seven in Test B. One person who reversed the time in one test is a stammerer. Two persons used the same time for the upward and downward work of Test A. Hence, of thirty persons, twenty-four added more rapidly upward than downward in Test A and twenty-three in Test B. Only one person was found who adds downward habitually and reversed both tests.

Time: The average time required to add *a single column* (forty figures) of Test A upward was 59.1 seconds, downward 61.3 seconds, or the average for each subject was 22 seconds more time required to add the downward than the upward half of the test, notwithstanding the records of four persons who added more rapidly downward are included in the averages.

The same averages for Test B were: upward, 58.3 seconds; downward, 60.2 seconds; difference, 1.9 seconds per single column, or adding the downward half of the test required 19 seconds longer for each person than adding the upward half. Here the records of seven persons who added a trifle more rapidly downward are included in the averages, and, of course, make them approach each other in amount.

It will be observed that inverting half the columns of Test B, so that the same associations appeared in the same order, saved .8 of a second on each single column of upward addition and 1.1 on each single column of downward addition. These differences probably represent the gain from adding a vestige of "immediate memory" to permanent memory. It seems evi-

[3]As stated later, one person was found who habitually adds downward.

dent that the subject adds 3, 6, 7, etc., more quickly from having added the same figures *in the same order* about twenty minutes before, but he does not add 3, 6, 7, etc., more quickly from having added 7, 6, 3 about twenty minutes before. I say a "vestige" of immediate memory, for after adding a column in one direction the subject had to add at least *eight* other columns before adding the same one in the reverse direction. While in general immediate memory demands that no impressions intervene to efface the immediate one, it appears from these slight but consistent gains that immediate impressions do not fade so rapidly that they cannot slightly shorten the time for forming the same associations (*i. e.,* combinations of the same data in the same order) a few minutes later. The differences in time are, it is true, of small magnitude, but they could be increased either by lengthening the columns or by using more than single-column addition for the test.

The average time for adding a single column is based on the addition of ten columns upward and ten downward in each test by each of thirty persons. Or the average in each test is derived from the addition of 600 single columns, 300 of which were added upward and 300 downward. But since the averages of *both tests agree that downward addition is the slower process,* that conclusion is based on the addition of 600 columns (23,400 separate associations) in each direction. Unless my group of thirty persons is an exceptional one, it is probable that the schools give considerably more practice in upward than in downward addition.

I have stated that the only objective difference between adding the upward and downward halves of Test B was a reversal of the habitual direction of the movements of eye and hand. These *motor* differences must, therefore, be the cause of the differences in time and accuracy in the two types of addition in this test. Test B involved only these motor factors, and Test A included these factors and a central one, namely, the different order of the associations to be made in the upward and downward adding (*i. e.,* 2, 8, 7, etc., upward; 7, 8, 2, etc., downward). Obviously, then, if we subtract the time required for Test B from that required for A, we shall have the time required for the central process of Test A, exactly as we com-

pute discrimination time by subtracting the time for the simple reaction to one signal from the longer time required to respond to a certain one of two signals.

The average difference per single column between adding upward and downward in Test B was 1.9 seconds; in Test A, 2.2 seconds. The difference of .3 of a second per single column in favor of Test B is due to the opposite orders of the combinations of digits to be made in the upward and downward additions of Test A. The difference is slight, but this is explicable on the ground that the average person has probably had quite as much practice in combining 2, 6, 7, etc., as in combining 7, 6, 2, etc. It seems probable, therefore, that so far as these two types of association are central, or unmodified by the immediately preceding perceptions, they require practically the same time, and that this is due to habit alone. At most the average difference in pure association time in the two cases was but 7 sigma. Hence, practically all the difference in time in favor of upward addition was due to peripheral causes, *i. e.*, to difference in the rates of perceiving the numbers.

Accuracy: So far I have discussed the time relations of upward and downward addition, but downward addition is the more accurate of the two processes. This leads to the inquiry, How does the single objective cause (Test B) operate to produce slower downward addition? There appears to be but one answer to the question: (1) The slower movements of eye and hand downward cause (2) a slower perception of the figures, and the associations are delayed merely by (3) having the data for combinations delivered to them more slowly. The figures of the tests were distinctly typewritten and in true vertical alignment. There were double vertical spaces between the lines. Hence, there could hardly have been either delay or inaccuracy in perception itself, but only in perception as depending on (waiting for) the preceding motor processes.

In Test A 54 per cent. of the total error was made in adding upward, 46 per cent. in adding downward. In Test B the averages were 51.8 per cent. and 48.2 per cent., respectively. Thus greater accuracy resulted from the slower (downward) addition in both tests, and while adding downward was more rapid in B than in A, it was also less accurate.

These results confirm the introspective belief that in the same person of average practice calculation is more accurate when it is performed more slowly. In other words, the results support the popular maxim, "Slow but sure." Since educational theory appeals nowadays to experiment rather than to experience, and is divided on the question of the appeal to voluntary attention, it seems possible that an experiment which shows that even momentary increase in effort of attention results in greater accuracy of associations is not without value.

Should it be found true that the perception of written figures is, in most persons, a more rapid process than associating them into partial sums, a condition which slows the rate of perception might make the associations clearer, more conscious and more accurate. Adding downward when one habitually adds upward would be one such condition. Voluntarily slowing the rate of perception to make the combinations more certain would be the same thing, except that it is subjectively caused. If, on the contrary, the figures could be combined into partial sums as rapidly as they could be perceived, very rapid addition should result in greater accuracy. This may be the case in persons who have had a great deal of practice in addition. But in either case greater accuracy would seem to result from having the perception of the figure and its association with the partial sum occur at the same moment. In our experiments with persons of average practice, therefore, it seems probable that the slower, downward addition was more accurate because it produced better co-ordination of perception with association.

In his investigation of counting and adding Arnett[4] found that "the cause of such errors as were made seemed to lie very frequently in the influence of some preceding figure still *delaying in the mind*. Burgerstein, in his studies of Vienna school children, found this to be one of the most common causes of error both in addition and multiplication." This language implies that each figure perceived arouses a mental duplicate of itself which may "delay in the mind." In persons practiced in addition this mental duplicate may not arise, but only the partial sum. Is it not possible that these two types of

images take the place of what, in addition to other data, is called the "anticipatory image?" And may not the difficulty of accurate addition lie in the fact that here the image can no longer be "anticipatory" nor "delayed," but must be accurately timed to meet the perceptions? Introspection tends to confirm this view, since one type of mental confusion is felt to be an experience of too many impressions coming at once.

The opinion that a habit of adding more slowly downward may have been formed because such adding is done to verify or prove the result of adding upward seems to be contradicted by the counting test. (See "Results of Test C," p. 93.)

Assuming now that there are two factors, habit and fluctuations of attention, involved in adding upward and downward, and that the fluctuations of attention work intermittently or momentarily, then if a majority of the subjects of the experiment have practiced adding upward more tha ndownward, the effect of habit should appear in five ways:

(1) A majority of the group should add the same columns more rapidly upward than downward.

(2) There should be fewer "reversals" than cases of the habitual time relation.

(3) But it should be possible to increase the proportion of reversals by decreasing the adding time, and to decrease the number by increasing the adding time. For a momentary, intermittent cause, like fluctuations of attention, must have greater relative effect if the adding time be momentary.

(4) The average time of the whole group of persons must be less for upward than for downward addition.

(5) Finally, greater accuracy should appear in the slower downward adding.

Since all five of these results appeared consistently in our experiments, the temptation lies near to claim that but two factors are present in adding, habit and attention, and that the former is more effective for speed, the latter for accuracy. Yet while the two factors are thus oppositely related, no effort of attention can obscure the effect of habit for more than five or six minutes at a time. Introspection agrees with this con-

[1]Arnett, L. D.: Counting and Adding. Amer. J. of Psych., Vol. 16, No. 1, p. 332.

clusion, and perhaps the reader will add that introspection reveals all the facts stated. I must reply that very few of my subjects had the slightest introspective knowledge of their habit of adding upward, and many of them overrated their ability to overcome the effect of habit by attention.

Results of Test C: If downward adding is in most persons both slower and more accurate than the same addition upward, what will be the effect of reversing our habit apparently acquired by reading from left to right? Each of the thirty persons who took Tests A and B was asked to take Test C. This consisted of counting "by ones" one hundred lines from left to right and an exactly similar series from right to left. While four persons used the same time in both directions, and two counted to the left a trifle more rapidly, the average individual time was 34.4 seconds when counting to the right, 37.1 seconds when counting to the left. The total error of all subjects when counting to the right was 62, when counting to the left 36. Hence, the reverse of the ordinary process is, in this case, as in addition, both slower and more accurate than the direct or usual process. The averages of all the tests point, therefore, to the same conclusion.

Arnett[5] says: "A greater degree of certainty was felt in counting from right to left than in the reverse direction, but no reason for this was discovered." The "reason" is doubtless the greater accuracy, and the introspective report is correct in this case.

Conclusions.

The agreement of the statistical results is stated on page 92.

1. Neural "paths" appear to be polarized by habit. So far as these experiments go the effect seems to be greater in peripheral paths than in central ones.

2. The effect due to the reading habit appears to be stronger than that due to the addition habit, since it is appreciable in a much briefer experiment. The average time for Test C was but 71.5 seconds.

[5]Counting and Adding. Amer. J. of Psych., Vol. 16, No. 3, 1905, p. 329.

3. Twenty-nine of our group of thirty persons gave evidence of an upward adding habit. One had formed the reverse habit.

4. For these persons of average practice a condition which increased the time even slightly resulted in greater accuracy in adding and counting.

5. There is some evidence to suggest that accuracy in addition depends upon the co-ordination of *equally conscious* or *equally mechanical* perceptual and associate processes; inaccuracy upon the co-ordination of a mechanical perceptual process with a conscious associative one.

6. Had a larger number of habitually downward-adding persons shared in the experiments, no correlation between time and the direction of adding would have appeared. Quite the same for accuracy and direction. Finally, accuracy and rate of adding have been found to stand in the relation of positive to negative. These facts help to show the extreme importance of Professor Judd's[6] demonstration that positive and negative cases often cancel each other, thus giving an appearance of no correlation where it can be shown to exist.

[6]The Relation of Special Training to General Intelligence. Educational Review, June, 1908, p. 28ff.

COMMUNICATIONS AND DISCUSSIONS.

A SUGGESTION FOR THE IMPROVEMENT AND EXTENSION OF MENTAL TESTS.

The present interest in mental[1] tests, particularly tests for deviations from the normal, means much for the future of education. There is one aspect of the mater to which I wish to direct attention. There is not in the Binet tests or any of the modifications yet proposed, sufficient system—no common plan running through the tests for the successive years. Prof. Terman's modifications and additions correct many of the most glaring defects of the Binet tests. There are some suggestions concerning the form of the tests and the method of stating the results obtained from them, which I wish to offer. These tests have been used to determine the number of years of retardation of the mind as a whole. Prof. Terman proposes, for instance, to take the child being examined as far as it can perform the tasks, then by adding up the tests successfully passed, determine the age of development represented by the child. Now, if we shift our attention from the mind as a whole to the mind as made up of several different functions, traits or capacities, it appears that the whole scheme of these tests and the statement of their results should be different. We should plan a series of tests for determining the degree of development of logical memory, rote memory, attention, imagination, association, and two aspects of mind more complex, learning capacity and reasoning. Even granting that the Binet tests may soon be made about perfect, they would still have to be changed very soon, for they are certain to come into a much more general use and their nature will soon be common property to school children. It seems to me that we need a national committee having authority to plan annually tests for the more important mental traits, and determine the normal performance for the successive ages. Take logical memory, for example, two short selections would be sufficient for all the different ages. But two or three supplementary tests should be prepared and standardized. In

the same way tests should be prepared and norms determined for attention, rote memory, imagination, etc. It is more important, it seems to me, to know specifically the condition of the child with reference to the development of the separate mental traits than to know his average performance with respect to them all. I am much in sympathy with Prof. Davenport[2] in his consideration of imbecility from the point of view of unit characters and heredity. What we wish to know is *whether and to what degree a specific mental trait is present.* To know that an eight-year-old child is two years backward does not mean much, but to know that he has poor logical memory, poor attention, or poor association, means much just as it means much to know specifically concerning the eyes, ears and other special organs.

I have in mind, also, a much more extended use of these mental tests than they have yet been put to.[3] From the time when a child enters school till he leaves it, accurate tests of his mental development should be made and careful records kept. The application of the Binet scale up to the present time has been chiefly to those children already picked out as abnormal, and probably in need of special training. If educational procedure is ever to be scientific, we must make accurate tests and keep the records, and this must be done for every child.

W. H. PYLE.
University of Missouri.

EDUCATION AT THE WASHINGTON MEETINGS.

The annual meetings of Section L, Education, were held at Washington Thursday and Friday, December 28th and 29th, as part of the sixty-third meeting of the American Association for the Advancement of Science. Thursday was devoted to general subjects—the new college entrance requirements in the morning, science teaching in the afternoon—while on Friday were given many short reports of scientific topics, or of investigations along educational lines. The Thursday morning meeting was a joint session with the American Federation of Teachers of the Mathematical and Natural Sciences, and on Friday afternoon the Section met with the American Psychological Association.

[1]For recent discussions see page 101 of this issue of the JOURNAL.
[2]C. B. Davenport. Heredity in Relation to Eugenics. New York: Henry Holt & Co. 1911. See also, G. M. Whipple, *Manual of Mental and Physical Tests*, 1910, p. 2, and my *Outlines of Educational Psychology*, 1911, p. 28.
[3]I have discussed this point and given a scheme for making the tests and keeping the records, in the *Outlines*, p. 236.

The reports of Prof. Jones for Columbia, Prof. Moore for Harvard, and Prof. Angell for Chicago on the changes in the entrance requirements were followed with close attention and many questions were asked of the speakers. Columbia proved the most conservative of the three, the real essence of the changes being that the university proposes to interpret its entrance examination grades in the light of the candidate's past record as supplied by the school. At Harvard the new scheme does not replace the old, but is simply an alternative plan whereby a good student from any school may be able to enter college although his school course has not been specifically modelled along Harvard lines. The entrance committee recognizes three factors in its decisions—the course of study pursued, the record made by the candidate, and the quality of his work as revealed by four examinations in standard subjects, which are designed to test, not the possession of specific information, but the quality of his mental powers. Any one of the three may be the deciding factor, each case being considered on its own merits.

The changes at Chicago proved the most radical of the three. By inspection from time to time, by records of the success or failure in college work of any school's graduates, and by frequent conferences and visits of teachers and principals to inspect college work, the university proposes to keep informed of the general quality of its accredited schools, and then to give them the greatest possible liberty in working out their own curricula to meet local needs. The requirements are general, not specific. Certain sequences of work must be carried out in order that the university may have a foundation upon which to build, the choice of these sequences must be from certain standard subjects, but within these limits the school is free, and the university is ready to supplement the preparation of its students wherever necessary by secondary courses in those subjects it considers essential to a university training.

The address of the retiring vice-president, A. Ross Hill, of the University of Missouri, emphasized anew the need for the development of the scientific spirit and the ability to solve real problems, rather than for the mere teaching of technical science. Prof. C. R. Mann, reporting on the work of a general course in physics at the University of Chicago, showed how this could be done. E. E. Slosson, associate editor of the *Independent,* called attention to the dearth of writers or lecturers able to deal with scientific subjects

in an interesting manner, and contrasted present conditions with those of the early days of science. He urged that courses in journalistic English be made an important part of scientific training.

Four papers were presented at the Friday morning session. Dr. J. E. W. Wallin reported an important investigation in Experimental Oral Orthogenics. A squad of 27 school children in Cleveland, Ohio, were given free dental treatment and hygienic instruction—filling of cavities, cleaning of gums, instruction in the care of teeth, 'fletcherizing' of food, etc.—and before, during, and from three to five months after this treatment were given a series of five mental tests to determine whether or not the remedy of physical defects had produced a corresponding increase in mental power. In spite of much individual variation, the results showed a decided gain in every test. The best proof of the benefit, and therefore of the importance of the work, is that, although all the members of the squad were laggards of from one to four years, only one failed of promotion in the term immediately following the treatment. The beneficial effects on the general health of the children was noticeable to children, parents, and teachers alike.

In the absence of M. B. Hillegas of Teachers' College, Prof. Torndike, chairman of the section, presented very briefly some account of the preparation of a scale for the measurement of merit in English writing. He read certain units of the scale and answered questions as to the use of the scale in practical work.

Prof. Norton's paper on "A Suggested Method of Experimentation in Education" was general in character, but it made very plain the fact that most educational problems are settled by custom, practical experience, opinion, etc., and that most of them are susceptible of strictly scientific treatment through control experiments within the school itself. On the other hand the purely scientific work done in the educational and psychological laboratories of the country is very rarely directly applied to actual school conditions. He urged, therefore, a closer co-operation between the scientific workers in educational psychology and the teachers and administrative officers of the school.

The last paper of the morning, quite by chance, came as a good illustration of the main point of the preceding paper, as it represented the scientific study of a real school problem by a worker directly concerned in its solution. S. A. Courtis, of the Detroit

HIGAN STATE NORMAL COLLEGE
TRAINING SCHOOL DEPARTMENT.

Home and Day School, gave a short report of the conclusions reached in an investigation to determine Standard Scores in Arithmetic. The results from the measurement of 9000 children in many schools revealed a condition of such great variability in all grades and schools examined that each grade, with respect to any single ability, was four or five grades wide. This result was shown to be remarkably constant, and to be indepedent of any errors in the method of securing the data. A comparative study of individual records disclosed great individual specialization, while the measurement of whole families at one time seemed to prove that the effect was due to inherited peculiarities of the individual mind. Mr. Courtis emphasized the variability of the individual as the most important factor in determining school procedure, and showed how, by supplementing class work by assignments of special work based upon the measured needs of the individual, it was possible to secure for each individual standard growth and standard scores at each grade.

Seven papers were presented at the joint session Friday afternoon, only three of which, from lack of space, can be reported in any detail here. Prof. Howard C. Warren, of Princeton University, described Montessori's Method of Teaching Writing and Reading. Interesting as was the paper, the criticisms of and comments on the method brought out many points usually passed over in such discussions, particularly the time cost of the method, and the possible effects on the health of the children in after life of such long hours of mental effort in early childhood. On the other hand, several who had visited the schools bore witness to the excellent results obtained and the apparent health and contentment of the children. The rapid and faddish introduction of the system into American schools was also commented upon unfavorably.

An interesting paper by Helen Thompson Woolley, of Cincinnati, described an attempted "Application of Experimental Psychology to the Problem of Vocational Guidance." The new child labor law of July, 1910, requires that working papers be issued only to children who are at least 14 years of age and have completed the 5th grade; and further that such children must report at the certificate office at each change of position. A special research bureau, working in co-operation with the Board of Education is attempting to make a careful study of individual children who go to work. The study includes the educational history, physical and psychological

measurements, home conditions, industrial history, and factory schedules of each child, and the measurements are repeated at each return to the certificate office. Standard psychological tests of mental traits are used under carefully standardized conditions. The accumulation of such data will in time show very plainly whether there is any connection between natural aptitudes in the traits measured and certain vocational pursuits.

Another paper of interest both to psychologists and educators was that by Dr. Starch of the University of Wisconsin, describing "An Objective Measure of Handwriting." By means of a graphometer consisting of a square of celluloid ruled in lines showing each degree of slant, and in horizontal lines above and below the base line, the mean variation in the slant of the first ten tall letters of any sample, and the mean variation of the letters above and below a base line, can readily be determined. From these two measurements a co-efficient of variability is derived, which, by reference to a fixed series of values, gives at once the quality of the handwriting so measured. Thorndike's scales, tested on this basis, yielded a series of values differing from each other under a constant law, and revealing the wonderful exactness of the units of the scale. It is to be hoped that Dr. Starch's work will be supplemented and confirmed by other investigators.

The remainder of the program* was devoted to investigations of learning and memory. Dr. Lough, of New York University, reported that the experimental study of the learning curves of many persons in the substitution test failed to show any characteristic "plateaus." Of the papers on memory, special interest attached to the report of Prof. Henmon, of the University of Wisconsin, that Ebbinghaus' well-known "law" as to the relation between amount to be learned and retention could not be confirmed. Henmon's results are in line with the recent work on economical learning and suggest strongly the need of retesting others of Ebbinghaus' classic conclusions, particularly since Ebbinghaus himself cautioned against generalization from results obtained on himself alone.

Detroit, Mich. S. A. COURTIS.

*The paper by L. W. Cole, on relative time and accuracy of adding upward and downward, which was read by title at this session, appears on page 83 of this issue. The paper by L. W. Kline, on the psychology of spelling, also read by title, and the papers by Professor Warren and by Dr. Starch will appear in forthcoming issues of the JOURNAL.—*Editors.*

ABSTRACTS AND REVIEWS.

RECENT LITERATLRE ON THE BINET TESTS.[1]

1. Alfred Binet et Th. Simon. *Le développement de l'intelligence chez les enfants.* Année Psychologique 14: 1908. 1-94.
2. O. Decroly et (Mlle.) J. Degand. *La mesure de l'intelligence chez des enfants normaux d'apres les tests de MM. Binet et Simon.* Archives de Psychologie, 9: No. 34, Jan. 1910. 81-108.
3. Alfred Binet. *Nouvelles recherches sur la mesure du niveau intellectuel chez les enfants d'école.* Année Psychologique, 17: 1911. 145-201.
4. Alfred Binet et Th. Simon. *La Mesure du developpement de l'intelligence chez les jeunes enfants.* Bulletin de la Société Libre pour l'Etude Psychologique de l'Enfant, Nos. 70 and 71, April, 1911, (11th year, No. 5.) 187-248.
5. Henry Herbert Goddard. *Two Thousand Children Measured by the Binet Measuring Scale of Intelligence.* Pedagogical Seminary, 18: June, 1911. 232-259.
6. Otto Bobertag. *Ueber Intelligenzprüfungen (nach der Methode von Binet und Simon.)* Zeitschrift für angewandte Psychologie, 5: No. 2, 1911. 105-203.
7. Katharine L. Johnston. *M. Binet's Method for the Measurement of Intelligence.—Some Results.* The Journal of Experimental Pedagogy and Training College Record, 1: No. 1, March 1, 1911. 24-31.
8. Alice Descoeudres. *Les tests de Binet et Simon et leur valeur scolaire.* Archives de Psychologie, 11: No. 44, Nov., 1911. 331-350.
9. Isabel Lawrence. *A Study of the Binet Definition Tests.* Psychological Clinic, 5: No. 7, Dec. 15, 1911. 207-216.

[1]For the sake of completeness, mention may also be made of an article by A. Ieronutti in the Rivista Pedagogica, 3: No. 3, 1909, and a monograph by Treves e Saffiotti, *La scala metrica dell' intelligenza.* Milano, 1911. These the writer has not been able to examine.

10. LEONARD P. AYRES. *The Binet-Simon Measuring Scale for Intelligence: Some Criticisms and Suggestions.* Psychological Clinic, 5: No. 6, Nov. 15, 1911. 187-196.
11. LEWIS M. TERMAN. *The Binet-Simon Scale for Measuring Intelligence: Impressions Gained by its Application on Four Hundred Non-selected Children.* Psychological Clinic, 5: No. 7, Dec. 15, 1911. 199-206.
12. J. E. WALLACE WALLIN. *A Practical Guide for the Administration of the Binet-Simon Scale for Measuring Intelligence.* Psychological Clinic, 5: No. 7, Dec. 15, 1911. 217-238.
13. EDMUND BURKE HUEY. *Backward and Feeble-Minded Children. Clinical Studies in the Psychology of Defectives, with a Syllabus for the Clinical Examination and Testing of Children.* Baltimore: Warwick & York, 1912. Pp. 189-202.

Perhaps no device pertaining to education has ever risen to such sudden prominence in public interest throughout the world as the Binet-Simon measuring scale of intelligence. Yet the scale itself was far from a sudden flash of genius in Binet's mind. Rather, as Professor Stern has recently said, it represents the culmination of a life work. As early as 1895 Binet had devised a series of tests which he offered as a tentative measure of "personality," and many of his publications during the succeeding decade give evidence of the continued search for an index of intelligence. In 1905 a practical demand, the determination of sub-normal pupils in French schools, led to the publication, in collaboration with Th. Simon, of a series of 30 tests of progressingly increasing difficulty, which should serve as an aid in ascertaining the intelligence of children. It was not until 1908, however, after these tests had been tried on 203 Parisian school children, that the authors hit upon their great idea—the idea which is now attracting so much attention—that of arranging the tests in a scale, each step of which should correspond to the normal intelligence of a given age between three and 13 years (1). In forming this scale some of the earlier tests were discarded and new ones were added so that the total was 56, with from three to eight tests for each age. It is upon this 1908 series that most of the recent investigations and discussions are based.

Among the first to make use of the scale were Decroly and Degand (2), who employed it in a private school in Brussels with 45 normal

children ranging from two years seven months to 12 years eight months. While they professed satisfaction with the general scheme, they criticised individual tests—some as too difficult or too easy for the age to which they were assigned, others as too mechanical, others as tests of specific training rather than of general intelligence.[2]

As a result of these and other criticisms and of further study and use of the scale, the authors published a revised edition of it (3 and 4) in 1911, in which the number of tests for each age was reduced to five (except for the fourth year), many of the objectionable ones were dropped, a few new ones were added, and the whole scale was rearranged to make it more representative of the performance of normal children. An important part of this rearrangement consisted in assigning many of the tests for older children to more advanced ages, so that whereas the 1908 scale gave tests for ages XI, XII, and XIII, the 1911 scale made provision for XII, XV, and adults. Since the 1911 revision was intended to supersede all other forms of the scale, it may be considered the definitive pronouncement of the authors upon the subject, and is here presented in tabular form, translated from the *Bulletin* (4). To render the table more serviceable, the position of each test in the 1908 series is indicated, and its age-rank is given according to the findings of the six most important applications of the tests to normal children.

In order to reduce the data of these investigations to a common basis, a test was considered as properly ranked for a given age when 75% of the children of that age were able to pass it.

In their earlier article (1) Binet and Simon gave no detailed account of the application of each individual test, and it was therefore impossible to determine what justification there was for the assignment of a test to a given age. Indeed, the lack of definite data in this article is so striking that one is led to wonder whether the scale was not first worked out in the study and later tested on children of various ages as opportunity permitted. In the subsequent paper (3), however, Binet gives a detailed table of his results —a table which he admits is compiled from both new and old material—and on the basis of this the ranking in column one has been determined. In the same article Binet recounts the results obtained by Levistre and Morlé with 20 children of each age from VII to

[2]For a detailed account of Decroly and Degand's criticisms, see Whipple's Manual of Mental and Physical Tests, Baltimore: 1910. Pp. 514-516.

BINET TESTS. 1911 SERIES.	1. Binet.	2. Levistre and Morlé.	3. Johnston.	4. Goddard.	5. Bobertag.	6. Terman and Childs.
1908.						
Same. III.						
1. Pointing to eyes, nose, mouth, etc.						
2. Memory for two digits						
3. Description of a picture (enumeration)						
4. Knowledge of family name						
5. Memory for sentences of six syllables						
Same. IV.						
1. Knowledge of own sex						III
2. Naming familiar objects						III
3. Memory for three digits						III
4. Comparison of length of two lines						III
V.						
V 1. Comparison of two weights				V	IV	III
V 2. Copying a square				VI	V	IV
New 3. Memory for sentences of ten syllables					V	
V 4. Counting four pennies				V	IV	IV
V 5. The divided rectangle				VI	VII	IV
VI.						
VI 1. Knowing morning and afternoon.	VII	VI		VI	VIII	V
VI 2. Definition of objects by use	VI	VI		VI	VI	IV
VII 3. Copying a diamond	VII	VII	VI	VII	VIII	VII
VII 4. Counting 13 pennies	VII	VII	VII	VI	VI	VII
VI 5. Elementary esthetic judgment	VII	VII		VI	VI	VII
VII.						
VI 1. Indicating right hand, left ear	VII	VII		VI	VII	VI
VII 2. Description of a picture (indicating relations)	VII	VII	VII	VII		IX
VI 3. Execution of a triple order	VII	VII		VI	V	IV
New 4. Indicating the amount of three 2-cent and three 1-cent stamps[3].	VIII	VIII	VI			
VIII 5. Naming four colors	VIII	VIII	VI	VII	VIII	VI
VIII.						
VIII 1. Comparing two objects from memory	VII	VIII	VI	VII	VIII	IX
VIII 2. Counting backwards from 20	VIII	VIII	IX	VIII	VIII	IX
VII 3. Unfinished pictures	VIII	VIII	VIII	VII	VII	VIII
IX 4. Knowing the date	VII	VIII	X	IX	X	IX
VII 5. Memory for five digits	IX	IX	VIII	VII	VII	IX

BINET TESTS, 1911 SERIES (Cont.).	1. Binet.	2. Levistre and Morlé.	3. Johnston.	4. Goddard.	5. Bobertag.	6. Terman and Childs.
1908. IX.						
IX 1. Making change (four cents from twenty)[4]	IX	IX	X	X	IX	XIII
IX 2. Definitions of familiar objects (superior to use)	XII	X	X	X		XIII
X 3. Naming nine pieces of money	IX	IX	XI	X	X	X
X 4. Reciting months of the year	IX	IX	IX	IX	X	IX
New 5. Replies to easy problem questions	IX	IX	X		IX	
X.						
IX 1. Arrangement of five weights	XII	XII	X	IX	X	XI
New 2. Copying two designs from memory		XII				
XI 3. Detecting absurd statements	XIII	XII	XIII	XI	XII	XII
X 4. Replies to difficult problem questions	XII	XII	XII	X	XII	XII
XI 5. Using three given words in two sentences	XII	XII		XI	X	XI
XII.						
New 1. Resistance to suggestion						
XI 2. Using three given words in one sentence	XII	XII	XIII	XI	XIII	XI
XI 3. Naming more than 60 words in three minutes	XIII	XV	XII	X	XIII	XV
XI 4. Definition of abstract terms	XV	XV	XIV	XII	XII	XIV
XI 5. Arranging words in a sentence	XIII	XII	XIII	XI	XII	XIV
XV.						
XII 1. Memory for seven digits	XV			XII	Ad.	XIV
XII 2. Rhymes	XIV			X	XV	XIV
XII 3. Memory for sentences of 26 syllables	XV			XV	XI	XV
New 4. Interpretation of a picture	XV					
XII 5. Problem questions	XV			XI		XIV
Adults.						
XIII 1. Drawing design from a cut in quarto-folded paper				XIII		Over XV
XIII 2. Juxtaposed triangles				XIII		XV
XIII 3. Distinctions between abstract terms				XIII		XV
New 4. Difference between king and president						
New 5. Substance of a prose passage						

[3]The French requirement is to count nine sous, three of which are double. Children are apt to be more familiar with coins than with postage stamps.

[4]The French have a single coin, the franc, for twenty sous.

XII (ranked in column two), and Miss Johnston's tests of varying numbers of Sheffield children, likewise from VII to XII years of age (column three). It should be noted that Binet's table gives the number of those who passed and of those who failed for each question at each age, that Levistre and Morlé record the number passing the test for every ten pupils examined, and that Miss Johnston merely indicates the total number of pupils of each age, and the number of those who fail on each question. The method of computing the rank of the test is therefore slightly different in each case. On the whole it would seem to the careful reader of this article that the statistical basis for the arrangement of the individual tests was not altogether convincing, and that furthermore the authors were not always guided in this arrangement by the statistics at their disposal.

The article in the *Bulletin* (4) presents the revised scale, gives detailed instructions for the conduct of each test, and indicates in very general terms the manner in which children of various ages respond to the tests. It is intended as a practical handbook for the administration of the tests, and is unquestionably the most authoritative guide for their use. There is no discussion of the merits of any test, as their validity in the present form is taken for granted.

The most extensive practical trial of the tests on school children has been made by Dr. Goddard (5), who superintended the examination of 1547 pupils in the Vineland, N. J., schools. Dr. Goddard is enthusiastic over the scale as an instrument for detecting mentally retarded children, and for revealing the extent of the retardation. Since the pupils show an almost normal curve of distribution about the "at age" group, he argues that the tests must be carefully and accurately graded age for age. To show the validity of the scale he has tabulated the responses of all pupils grading "at age," giving the number of those who passed and those who failed on each question. From this table the ranking in column four has been computed. It is unfortunate that with this mass of valuable material in hand no table is given to show the responses of all the pupils examined to each question. While it may seem fairer to the scale to judge of the arrangement of the questions by the reactions of those pupils grading "at age," it must be borne in mind that they number only 554, and thus constitute but a little over one-third of the 1547 children tested. In an inquiry as to the validity of the scale, however,

one must ask what percentage of children of a given age *taken at random* are able to answer each question satisfactorily, and this demand would be more nearly met by tabulating all the results than by taking only those from a selected group.

Dr. Bobertag (6) gives the most thorough psychological analysis of the individual tests that has yet been made. He examined over 800 children in the Breslau schools, and while he presents no formal tables, he indicates in the discussion of each question the percentage of pupils of a given age that gave satisfactory answers. From these statistical notes the rankings in column five were determined. The rankings in column six were estimated from the table of Terman and Childs on page 72 of this issue.

An inspection of the table shows that in spite of many differences in the rankings of individual questions there is a surprising agreement in the results of the different investigators. This is the more striking in view of the widely separated communities in which the tests were made. France, England, Germany, New Jersey and California are widely enough distributed to be fairly representative of the civilized world. From this agreement we may draw two conclusions: First, that a large number of the tests are correctly placed; second, that the following tests should be shifted: VI, 3, 4 and 5 to VII; VII, 3 to VI; VII, 5 to VIII; VIII, 4 to IX; IX, 1 and 2 to X; X, 3, 4 and 5 to XI or XII; XII, 2, 3 and 4 to XIII or XIV. It is noteworthy that many of these shifts would return the tests to the same places they held in the 1908 series. Whether the fact that these results were obtained with the use of the 1908 scale has any bearing on the ranking of the tests, only the use of the 1911 scale under the same conditions could determine. We are informed that Dr. Goddard is at present conducting such a check examination in Vineland.

A large part of Dr. Goddard's paper (5) is devoted to a detailed comparison of the mental age of pupils, as determined by the Binet scale, with their grade in school, or their pedagogical age. He finds that only about half of the mentally "over age" and "under age" pupils show a corresponding advancement or retardation in their grading. Half of them are therefore said to be improperly graded. From his general experience the writer suspects that this is probably the case, yet it does not seem to follow from a consideration of the Binet tests alone. A pupil's grade in school should depend on his

mastery of school subjects. The Binet tests are supposed to be tests of *native intelligence,* and to neglect as far as possible all that has been learned in school. Hence it might well happen that a natively bright child should properly be retarded in his school standing, and a natively dull child by persistent application might well be pedagogically advanced. It is quite beyond the mark to carry these tests directly over to school work and make them a basis or a criterion of school grading. For this we need other tests worked out from the pedagogical point of view on the basis of school subjects.

In the most recent general study of the Binet tests (8) Miss Descoeudres selected one bright and one dull boy, and one bright and one dull girl from each of the six primary grades of the Geneva schools—making 24 pupils in all. Each of these was tested with all the questions of the 1908 scale. The results are not tabulated nor are they arranged in convenient form for tabulation. The author finds the scale too easy at the bottom and too hard at the top (as does Terman in the present issue). The bright pupils do better than the dull ones (57% to 43%), and the boys make a slightly better showing than the girls (53% to 47%). On the whole there seems to be some correlation between the teachers' estimates of "bright" and "dull" and the degree of intelligence indicated by the Binet tests.

Miss Johnston (7) recently tested 218 girls ranging from six to 15 years of age. She also finds the tests for VI and VII too easy, while those for X to XV are too difficult. Tests that gave particular difficulty were IX, 1, 2 and 3; X, 1, 4 and 5; and all of those in XII. It is impossible to obtain from her account any clear picture of the rank which should be given to individual tests.

Dr. Bobertag (6) has examined each question of the 1908 series in detail from a psychological point of view, showing the kind of mental activity involved in the test, giving minute directions to ensure clearness and accuracy of response from the child, rejecting some of the tests as having little value for intelligence, commending and amplifying others. X, 1, the arrangement of five weights, he considers an excellent test because it does not depend in the least upon learning, nor does it involve a verbal response. VII, 2, description of a picture, is a good test if the picture is properly selected. The definition tests, VI, 2, and IX, 2, he finds among the most interesting of the list, and yet the most difficult to evaluate, for in

many cases it is doubtful whether a definition by use is not superior to any other definition that might be given.

The definition test VI, 2, together with VIII, 1, comparing objects from memory, XII, 4, definition of abstract terms, and adults, 3, distinctions between abstract terms, has been employed as a written school exercise by Miss Lawrence (9). She had 784 pupils tested at the same hour in the day. She, too, found that a large number of pupils tested above age in the lower grades, but that an almost equally large number were below age in the higher. She thus confirms the statements of Terman and Descoeudres that the scale is too easy at the lower end and too difficult at the upper.

But these are not the only criticisms that are leveled at the scale. Dr. Ayres (10) complains that it tests ability to use words rather than to perform acts. This is true in a high degree, for 43 of the 54 tests are verbal. Yet is it doubtful whether this is a very serious drawback, for our ordinary judgments of intelligence are based largely on linguistic ability. Again, many of the tests depend upon recent environmental experience, such as those involving time and money, or upon the ability to read and write, which is a matter of schooling. Further, many of the tests involve the immediate memory of words and numbers, but almost all investigators of memory have found that there is no correlation between memory span and intelligence. Finally there are too many "Puzzle tests," and an unreasonable emphasis is laid on the definition of abstract terms. To this it may be replied that the ability to cope with new situations is one of the most important criteria of practical intelligence, and there is perhaps no better way of testing this than by judiciously selected puzzle tests. Furthermore, the chief mark of the superiority of human intelligence over that of animals is man's ability to use abstractions. Tests which would reveal this ability in its different stages would have a high value as indices of intelligence.

Professor Terman (11) points out the fact that the value of Binet's work lies chiefly in the demonstration of the feasibility of such scales and their utility for particular purposes. Heretofore tests of intelligence have been haphazard and without any central controlling principle. This principle Binet supplied by correlating inteligence with age. While the Binet scale is far too limited in extent, far from accurate, and requires improvement in many re-

spects, it deserves the credit of a pioneer work. The author emphasizes the necessity of connecting tests of intelligence with tests of physiological age.

Dr. Wallin (12) notes that the general use of the Binet tests will soon yield a mass of valuable data about children if it is collected uniformly. To this end he furnishes detailed and carefully worded directions for giving the 1908 tests, insisting that this series needs to be tried out more carefully before further revisions are attempted. The same purpose has animated Dr. Huey to insert in his book (13) a syllabus for the use of the 1911 Binet scale with certain modifications suggested by Goddard.

In all of these studies of the Binet-Simon measuring scale of intelligence the writer has been impressed by the fact that not a single investigator has raised the question "What is native intelligence? What does the term signify? What ground have we for thinking that these questions will reveal what we mean by intelligence?" That some of them refer to only a very limited sort of intelligence is fairly obvious. Their selection or rejection seems to have been largely a matter of caprice thus far. Is it not time to begin at the other end, to attempt an analysis of the complex, popular term intelligence, and on the basis of this analysis to devise tests which will correlate with chronological and physiological age? Such a task would be worthy of the best efforts of any experimentalist.

J. C. B.

HENRY H. GODDARD. *Four Hundred Feeble-Minded Children Classified by the Binet Method.* Pedagogical Seminary, 17: September, 1910, 387-397.

The application of the Binet-Simon tests to children in the Vineland, N. J., institution afforded a classification of the 400 individuals tested that was felt to be satisfactory and reliable, when checked in various ways by reference to the experience of the institution and to the opinions of those having the children in charge. The children ranged in mental age from one to twelve years.

One source of difficulty is of special interest. Suppose a feeble-minded person to achieve the mental development corresponding to the Binet-Simon ten-year grade, but suppose him actually to live to be 50 years old: what effect will the long experience have upon his performance? Speaking generally, such an individual will have learned to do a great many things more than any single ten-year-old

child does: his superior performance, then, will indicate longer experience, not greater intellectual maturity. Cases so extreme as the one just assumed must be thrown out of consideration in estimating the correspondence between the diagnosis of the tests and the diagnosis of the institutional force.

Speed of performance with the "form-board test" appears to correlate very well with the psychological age as indicated by the Binet-Simon tests.

Those shown by the mental tests to rate psychologically as one or two years may be classed as idiots, those rating three to seven years as imbeciles of low, middle, and high grade, those rating eight to twelve years as feeble-minded (in the narrow usage of the term) —the last two years being high grade feeble-minded, or *morons* (to use the term proposed here and now generally adopted).

The problem of the "moral imbecile" is raised by Dr. Goddard, who proposes the novel explanation that this type represents a nine-year old arrest of development—the idea being that lying, thieving, and the like are instinctive tendencies normally nascent at nine, and overcome thereafter in the ordinary child. Had the arrest come earlier, the imbecile would not have developed these instinctive tendencies: had it come later, he would have developed sufficient reasoning power to overcome them.

The reviewer is not disposed to dispute this hypothesis, though he would like to see further evidence before accepting it: he can not, however, see the force of Dr. Goddard's assertion that in primitive times, "lying and stealing were virtues and *consequently* became instinctive." This looks pretty much like the inheritance of acquired characteristics.

CORNELIUS HOOD.

Cornell University.

EDMUND B. HUEY, PH.D. *Retardation and the Mental Examination of Retarded Children.* (Reprint from Journal of Psycho-Asthenics, Vol. XV.)

The writer of this article sets forth a point of view which, though not novel, has had, in my opinion, far less attention than it deserves. He calls attention to the fact that the Binet-Simon diagnostic tests of mental development are essentially tests of intelligence; that they stop short at the pubertal period, and that we have as yet no satisfactory extension of the series into the highly important period of adolescence. He reminds us that feeble-mindedness, as encount-

ered in institutional cases, begins at about the 8th psychological year, but stops short at the dead-line of the 12-year intelligence (only two in an institution of 1300 reached the Binet 13-year level). Now, may we assert that any person who can pass the 13-year tests may be considered normal in intelligence? Perhaps, yes, if by "normal intelligence" we mean ability to "get on" outside a protective institution. But to get on as a laborer does not mean to get on as a physician or as an attorney. Every professional and social class has its dead-line of minimal ability, and in the activities of adult life individuals become sorted out into strata according to their capacities—capacities conditioned not merely by intelligence, but also by "normality of will, of self-assertion and self criticism, of social sense and attitude, and of emotional control. And here, in this rich, but precarious field of functionings are found the next higher rungs of the retardation ladder."

We need, then, according to Huey, adequate tests to detect the incapacity for making these higher adaptations of mental development, tests which would naturally apply to the period following the Binet-Simon scale, since adolescence is the period when these adaptations are or should be perfected.

Most of the chronic psychoses and neuroses classed as neurasthenia, psychasthenia, and hysteria are cases of retardation. "Retardations occur continuously on up the years of growth to maturity." And these "higher level retardations should especially concern us, since they cripple the individuals who have a part to play in society." Defectives above the 12-year level "are presenting the most serious of the problems of society."

The tests to be desired must deal with control of movement, attention, capacity for synthesis, emotivity, feeling, the play and art interests, learning, memory, and ability to report, reasoning, the formation of ideals, temperament, and especially the "self-estimations and self-relationships." CORNELIUS HOOD.

Cornell University.

ALICE DESCOEUDRES. *Exploration de quelques tests d'intelligence chez des enfants anormaux et arriérés.* Archives de Psychologie, 11: November, 1911, 351-375.

The author, at the suggestion of Claparède, sought to determine by trial the merits of a set of 15 tests of intelligence, six taken from

the Binet-Simon series, and nine others. The plan was to select 14 backward and abnormal children of both sexes, and ranging from six to 14 years of age (chronologically), to determine by general observation of the daily work of these children their rank in intelligence and to compare, by correlational statistics, their rank in each of the 15 tests with this originally estimated rank.

The tests taken from the Binet series were (1) description of pictures, (2) definitions of familiar objects, (3) differences between familiar objects recalled in memory, (4) naming words for three minutes, (5) replies to four problem-questions, and (6) naming four coins. The tests added by the author were: (7) stringing 20 beads as rapidly as possible, (8) sorting into four boxes by touch alone five each of coffee beans, kidney beans, peas, and grains of wheat, (9) test of "imagination" (inventing a speech for children or animals shown in three pictures), (10) test of visual imagination (naming six incomplete drawings), (11) reconstructing simple puzzle pictures (a horse and a dog, each cut into six parts), (12) cancellation, (13) computation (six very elementary problems), (14-15) a group of five and one of nine objects is shown (or named) to be reproduced by drawing (attention and memory).

The results are striking. When the rank of each pupil in each test is summated, the resulting total rank shows a correlation with the originally estimated rank of intelligence, of $r=.991$, *p.e.* .001. When the tests are considered separately, the correlation ranges from $r=.878$ to $r=.509$. In order of height of correlation with estimated intelligence, the several tests assume the following order (using the numbers given above): 3, 13, 1, 5, 8, 2, 7, 9, 11, 6, 12, 14, 10, 15, 4. The first six of these (differences between objects, computation, description of pictures, problem-questions, sorting seeds, and definitions) correlate with estimated intelligence by .80 or more.

If these results, secured with a small number of backward children, can be confirmed by more extensive investigation with both backward and normal children, they will go far to establish the reliability of simple psychological tests in the diagnosis of degree of intelligence, and they will suggest a useful method of testing tests themselves in terms of their value as indicators of mental age.

Cornell University. CORNELIUS HOOD.

GUY MONTROSE WHIPPLE. *Relative Efficiency of Phonetic Alphabets.* Educational Psychology Monographs. Baltimore: Warwick & York, Inc., 1911. Pp. vii, 52. Paper, 35 cents.

More and more the methods of experimental pedagogy are called upon to throw much needed light upon practical questions. The subtitle of this Monograph, "An Experimental Investigation of the Comparative Merits of the Webster Key Alphabet and the Proposed Key Alphabet Submitted to the National Education Association," indicates its scope. The most extravagant claims have been made by a small group of phonetic enthusiasts and educational reformers regarding the psychological and pedagogical advantages of the Proposed Key Alphabet. Professor Whipple has subjected some of these claims to the "acid test" of experiment, and has found them wanting. The questions raised were the following: Is the Proposed Key more easily or less easily learned than a key like that now in common use, the Webster Key? When once learned, is it remembered as well as the Webster Key? Can it be used as a guide to pronunciation with as great facility and accuracy as the Webster Key?

Four sets of experiments were undertaken: (1) On beginners in reading; (2) on fourth-grade pupils; (3) on seventh and eighth-grade pupils who had had no previous training in phonics or in the pronunciation system of any dictionary, and (4) on college students. The method of the experiments, the materials used and the results obtained are set forth in the Monograph with scrupulous regard for the details of scientific procedure. One group of pupils was taught the Proposed Key, another the ordinary key, and after the same number of learning lessons it was found that from 47.3 per cent. to 71.3 per cent. more errors were made with the Proposed Key. One week after learning had ceased 55.6 per cent. more errors were made with the Proposed Key. The learning of the Proposed Key by college students exacted 41 per cent. more repetitions, with 85 per cent. more errors, than the learning of the Webster Key. In conclusion, the author summarizes his results thus: "Careful experiments with children in the public schools and with college students show that, without exception, the Proposed Key is more difficult to learn, and that, when once learned, it is more slowly and less accurately used, whether we measure the facility of its use by tests of pronunciation of the isolated symbols, pronunciation of proper names, pronunciation of nonsense syllables, or by tests of phonetic transcription of simple words."

J. C. B.

NOTES AND NEWS.

THE "KEY ALPHABET" AND ITS ADVOCATES.

The publication (as a number of the *Educational Psychology Monographs*) of Professor Whipple's tests of the Phonetic Key Alphabet has brought forth a veritable little storm of abuse and vituperation. It will be remembered that this Alphabet was proposed for the indorsement of the Department of Superintendence at the Mobile meeting with the distinct and emphatic assertion that it was *easier to learn, easier to remember,* and *easier to apply* than the system of diacritical markings now in common use. It is clear that these assumed virtues are distinctly and exclusively psychological in their nature; the committee which constructed the new alphabet was made up exclusively of philologists and school administrators. They put forth their joint production under an untested psychological assumption. Professor Whipple put the Alphabet to the test of actual practice under controlled conditions. His results did not substantiate the statements made by the committee.

Now comes President H. H. Seerley of the Iowa Teachers' College, a member of the committee, with a circular-letter to the members of the Department of Superintendence, citing the opinion of five philologists against the validity of Whipple's conclusions. One sentence in Mr. Seerley's letter may be taken as indicative of his position in the matter:

> "This Key is a free key that anyone may use or anyone can reject, and there is no good reason why so much disturbance should be produced by those who do not care to use it."

With due respect to President Seerley, we maintain that, when an attempt is made to foist upon the public schools through the leading association of public-school workers a theory that rests *by the statement of its advocates* upon untested psychological assumptions, it is not only the right, *it is the duty,* of the educational psychologist to determine in how far the assumptions can be justified. If the Key had been proposed on the ground of its philological superiority, it

would not be the province of the specialist in educational psychology to pass judgment. But it was not urged from this point of view.

In the name of our profession and in the name of some twenty million school children innocent of both philology and psychology, we protest against the uninformed and undignified vituperation to which Professor Whipple has been subjected. He was clearly within his rights in making the investigation; his published conclusions were justified by his published data; his statements were good-tempered, impersonal and guarded. His tests may be repeated. His methods may be improved; his conclusions may be reversed; but the fact still remains that he was the first to subject the claims of the new Alphabet to the test of actual practice. He was the first to adopt a simple expedient to determine whether the Alphabet really "worked" upon a limited number of individuals before he should endorse it for universal adoption. W. C. B.

SEEING IS BELIEVING.

We had thought that the teachings of Bacon, Darwin, Huxley and others of that brilliant group whose labors meant so much for science had colored indelibly the intellectual life of the present day. In our simple faith we forgot for the moment that intellectual habits and attitudes often fail to "transfer." Rude is the awakening. Among the opinions cited by President Seerley in his circular-letter is the following from Professor C. H. Grandgent, chairman of the Romance Department, Harvard University:

"His [Whipple's] results are contrary to reason. One has only to look at the two alphabets to see which is the shorter, the simpler, the easier to write, and the better adapted to mnemonic processes."

Shades of Giordano Bruno! One has only to look at the sun to see that it wheels from east to west above a stationary earth! One has only to look about one to see that this world of ours is flat!

W. C. B.

SOME GRATUITOUS INSINUATIONS.

The Secretary of the Simplified Spelling Board, contributing to President Seerley's symposium, makes an interesting suggestion: "Another fact bearing upon the matter might be brought out by inquiring whether Professor Whipple undertook this investigation of his own motion, or whether he was asked to undertake it, for a consideration, by some party in-

terested. It would also be highly interesting to know the name of that party. Perhaps a clue to the name might be found in the fact that the investigation was undertaken soon after the February meeting, and was ended in time for the July meeting, of the National Education Association. A further line of inquiry might be opened with the question why that party did not publish the report, but left it to be published by a concern in another city, bearing the romantic name of 'Warwick & York.'"

We fail to see the pertinence of either question. The propositions laid down by the committee in advocating the indorsement of the Alphabet by the Department of Superintendence made extravagant and untested assertions regarding the psychological merits of the Alphabet. Professor Whipple found that, in typical instances, these statements were not substantiated. If the advocates of the Alphabet are dissatisfied with the tests, it is their privilege to repeat them. Repeated tests will inevitably disclose any factors that might have operated to render Whipple's conclusions untrustworthy. Until such tests are made, the insinuations in the above citation constitute a purely gratuitous and entirely unjustified attack upon the integrity both of the investigator and of two reputable publishing houses. It is one thing to infer that investigations have been subsidized; it is another thing to make the double inference that the subsidy has colored the results. Until the latter has been proved, the former is of no significance. Truth is truth, whether it is paid for or not.

W. C. B.

THE KEY ALPHABET AS THE TRIUMPH OF THE CENTURY.

And speaking of the way in which publishing houses are concerned in this matter, the following citation from a circular widely distributed by a publishing house that seems to be exploiting the new Key can hardly be omitted from the discussion:

"The final completion and adoption of this Alphabet by the Department of Superintendence, it is believed, will pass into history as one of the greatest, if not *the* greatest, of the educational triumphs of the English-speaking world during the century."

When it is remembered that this circular was distributed some months *after* the results of Whipple's investigations had been made known at the meeting of the N. E. A. at San Francisco; and when it

is remembered that the circular neither referred to Whipple's conclusions nor reported factual evidence controverting them, this characterization of the new Key as one of the greatest educational triumphs of the century strikes us as about the most flamboyant bit of advertising that we have ever known educational publishers to indulge in. We suggest that the Secretary of the Simplified Spelling Board look carefully to his own backyard before he makes gratuitous references to the publishing house "bearing the romantic name of 'Warwick & York.'" W. C. B.

THE VIEW OF A PHONETICIAN.

In connection with the discussion of phonetic keys, the following statement by the well-known phonetician and educator, Professor Alexis F. Lange, Head of the Department of Education of the University of California, may be of interest to our readers. The statement was made to the writer for public use at the San Francisco meeting of the N. E. A.:

> "As a student of phonetics and a teacher of English for over twenty years, I am opposed to the adoption of the report [on the Proposed Key] for the following chief reasons:
>
> "1. The proposed phonetic key is not sufficiently scientific to serve the pupil, either now or later on, for scholarly purposes.
>
> "2. It is not simple enough, nor pedagogical enough, to really assist the pupil in learning English.
>
> "3. I seriously doubt whether any system will be satisfactory for purely educational purposes that goes outside of English letters and their traditional values in order to designate English sounds.
>
> "4. I seriously question the propriety of committing the N. E. A. to a plan of which only experts are able to judge adequately." J. C. B.

THE AMATEUR AND THE BINET-SIMON TESTS.

We have endeavored to keep our readers in touch with the development of the Binet-Simon scale for diagnosing mental age, and propose to publish several critical appraisements of the scale during the next few months. It may not be out of place, however, to comment here upon the popularizing of the method and upon some of the consequences that may flow therefrom.

The Binet scale, after several years of trial by its originators and by various educational psychologists, is now evidently approaching a period of extensive use in the public schools. We understand, for instance, that New Jersey has enacted a law which directs a mental examination to be made of all school children who are pedagogically retarded. It is fair to assume that the Binet scale will be used: but who will conduct the examinations? Our large city school systems, with a few conspicuous exceptions, have not yet realized the imperative need of establishing psychological clinics or bureaus of child study, or even of employing a consulting psychologist. When psychological examinations come to be introduced, they tend, failing a psychologist, to fall into the hands of medical inspectors, which is bad enough, or to be assigned outright to classroom teachers, which, perhaps, is still worse. We have no quarrel with the use of the scale in the public school: properly used, it is of direct and practical value; but improperly used, it will become a farce which can but bring discredit upon psychology and retard the movement for its application to educational practice.

There is nothing about the conduct of the Binet-Simon tests that is intrinsically difficult, yet there is a source of error inherent in the use of any psychological procedure, which, as experience shows, is surmountable only by drill in psychological experimentation. I refer to the difficulty of following directions. No one who has drilled students in the laboratory has failed to be struck with the impossibility of laying down fool-proof directions for the conduct by an amateur of a psychological test. The motto: "Don't monkey with the method" ought to be hung on the walls of every laboratory. To cite examples: a young examiner is reported to me who "invented" a new set of test-questions to supplant a portion of the original Binet series, then varied these questions from child to child, and finally wondered why the Binet method was so unreliable. Another substituted letters for digits in the memory-span test because he "liked it better that way."

We hope the Binet scale will succeed in the public school: if we may not hope for psychologists to apply it, may we not at least hope that the school authorities will send their examiners to some institution where they may receive instruction in general and educational psychology, and special drill in the handling of what is really an exceedingly delicate psychological instrument? G. M. W.

THE PRESENT TREND IN PSYCHOLOGY.

The trend of psychological interest is nowhere better gauged in this country than at the annual meeting of the American Psychological Association. The meeting just held in Washington was characterized by a dominating interest in *applied* psychology. One joint session was held with Section L, Education. One session was devoted to psychology in medicine, and proved exceedingly stimulating and fruitful. Among the physicians who took part were Doctors Franz, Adolf Meyer, Southard, Morton Prince, Jeliffe, Hoch, Kober, and Allen Starr. This session proved a sort of professional love feast in which psychologists and physicians exchanged a very hearty "we need you." About one-half of the remainder of the program dealt with phases of applied psychology. This tendency was also voiced by the president's annual address, which surveyed the scope and significance of mental measurement as a foundation for the applied psychological sciences. There was not a single paper representing the philosophical attitude. The absence of members of the Philosophical Association was regrettable. It is to be hoped that the Philosophical Association may change its plan of meeting in isolation, for it is perfectly clear that the psychologists must meet where they may have joint sessions and come into contact with the educationists, the physiologists, the physicians, and other related bodies, but psychology will ever remain bound in close ties with philosophy. C. E. S.